

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 11, 2003, 18:07:07 : Search time 251 Seconds
(without alignments)
10371.753 Million cell updates/sec

Title: US-09-805-694B-1

Perfect score: 1156

Sequence: 1 gcgcgcgcattggttcctt.....tctcctcttgagcgcgcgc 1156

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_101002:*

1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1156	100.0	1156	22	AAD17527
2	1156	100.0	2970	22	AAD17528
3	167	14.4	1807	18	AAT74041
4	154.4	13.4	1056	18	AAT74043
5	154.4	13.4	1392	24	ABN83213
6	154.4	13.4	1849	18	AAT74042
7	152.8	13.2	1056	18	AAT79546
8	151.2	13.1	1577	18	AAT90169
9	147	12.7	1553	18	AAT90170

10	136.6	11.8	1083	21	AAC4888
11	135.4	11.7	1538	21	AAC46206
12	134.2	11.6	1062	22	AAT44804
13	132.8	11.5	1600	22	AAT44793
14	130.8	11.3	1698	21	AAC40001
15	128.6	11.1	1369	21	AAC51084
16	128.6	11.1	2167	22	AAC44797
17	126	10.9	1160	21	AAC33767
18	125	10.8	1201	21	AAC41593
19	121.8	10.5	1112	21	AAC42369
20	120.2	10.4	1434	18	AAT90172
21	120	10.4	910	19	AAV66759
22	118.6	10.3	1390	18	AAT90171
23	118.6	10.3	1441	18	AAV90173
24	117.2	10.1	765	22	AAT44805
25	117	10.1	1186	16	AAO94588
26	117	10.1	1186	18	AAT65477
27	117	10.1	1186	19	AAV36970
28	117	10.1	1186	21	AAZ35693
29	117	10.1	1186	22	AAZ47380
30	115.4	10.0	1186	19	AAV27947
31	114.8	9.9	1474	18	AAT90174
32	113	9.8	1370	21	AAC32968
33	109	9.4	1645	20	AAV82458
34	108.2	9.4	1879	21	AAC44552
35	103.8	9.0	2446	21	AAC46450
36	99	8.6	1224	21	AAC40331
37	98.8	8.5	1474	22	AAT44796
38	97.6	8.4	883	22	AAE44791
39	97.6	8.4	948	22	AAE44790
40	95.8	8.3	1077	21	AAC51435
41	95.8	8.3	1282	24	ABN98252
42	93.2	8.1	1305	21	AAC47882
43	92.6	8.0	1407	22	AAE44802
44	92.6	8.0	1463	22	AAE44787
45	92.4	8.0	785	21	AAC33494

ALIGNMENTS

RESULT 1
AAD17527
ID AAD17527 standard; DNA: 1156 BP.
XX
AC AAD17527:
XX
XX 10-DEC-2001 (first entry)
DT
DE Soybean vacuolar protein (P34) encoding DNA.
XX
XX Soybean; allergen; transgenic plant; p34 protein; Gly m Bd 30K; Gly m 1;
KW soybean vacuolar protein; Gly m 1A; Gly m 1B; rGly m3; Glycinin G1;
KW alabab; food; infant formula; animal feed; coating; salad oil; syrup;
KW spraying oil; roasting oil; frying oil; cracker; confectionery product;
KW snack food; topping; sauce; batter; bread; baking mix; dough;
KW ds.
XX
OS Glycine max.
XX
XX
XX WO200168887-A2.
PN
XX
XX 20-SEP-2001.
PD
XX
XX 15-MAR-2001; 2001MO-US08254.
PF
XX 16-MAR-2000; 2000US-0189823.
PR
XX
XX
XX (DUPO) DU PONT DE NEMOURS & CO E. I.
XX (PION-) PIONEER HI-BRED INT INC.
XX Jung R, Kinney AJ;
XX

xx		Recombinant expression construct to lower allergen (e.g. Gly m Bd 30k)
pt		content of a soybean, comprises a nucleic acid fragment encoding the
pr		allergen, useful for producing soybean plants which can be used to make
ps		soybean products
xx		
ps		Claim 4; Page 48-49; 57pp; English.
xx		
cc		The patent discloses hypoallergenic transgenic soybeans and recombinant
cc		expression constructs to lower soybean vacuolar protein, commonly known
cc		as P34 (Gly m Bd 30K or Gly m 1) and other allergens such as Gly m IA,
cc		Gly m IB, rlgly m3 or Glycinin G1 (alabibi). The allergen content of the
cc		soybean is reduced by sense suppression which is accomplished by using
cc		the expression construct that comprises a nucleic acid fragment encoding
cc		the allergen. The constructs are useful for producing hypoallergenic
cc		transgenic soybean plants which can be used to make hypoallergenic
cc		soybean products which can be used in a variety of food (e.g., infant
cc		formulas) and animal feed applications. The oil made from seeds of the
cc		hypoallergenic transgenic soybean plants can be used as ingredients,
cc		as coatings, as salad oils, as spraying oils, as roasting oils, and
cc		as frying oils. The foods in which the oil may be used include crackers
cc		and snack foods, confectionery products, syrups and toppings, sauces,
cc		batter and bread mixtures, baking mixes and doughs. The present
cc		sequence is a DNA encoding Kunitz soybean trypsin inhibitor (KSTI
cc		or Kt13), a minor soybean seed allergen.
xx		
SQ		Sequence 2970 BP; 1018 A; 547 C; 527 G; 878 T; 0 other;
		Query Match 100.0%; Score 1156; DB 22; Length 2970;
		Best Local Similarity 100.0%; Pred. No. 0;
		Matches 1156; Conservative 0; Mismatches 0; Indels 0; Gaps 0
OY	1	GCGCCGCATGGGTTCCTTGTTGTTGCTTCTTTTCCCTCTTAGGTCTCTCTAGTT 60
Dd	622	GCGCCGCATGGGTTCCTTGTTGTTGCTTCTTTCCCTCTTAGGTCTCTCTAGTT 681
OY	61	CCAGCATATCAACTATGTTTCATATTGACCTTGACCTTAACAAGATTACCACACAGA 120
Dd	682	CCAGCATATCAACTATGTTTCATATTGAGACCTTGACCTTAACAAGATTACCACACAGA 741
OY	121	AACAGTGCTTCATCGTTCCAACTATGGAAGAGAGCAGTACGAGTGTCTACCAATACC 180
Dd	742	AACAGTGCTTCATCGTTTCACATGTTCCACATGGAAGAGAGCAGTGTCTACCAATACC 801
OY	181	ACGAAGAAGAGCAAGAAGAGACTTGGATTTCATAAATTAATCTGAACTATCACAGGACA 240
Dd	802	ACGAAGAAGAGCAAGAAGAGACTTGGATTTCATAAATTAATCTGAACTATCACAGGACA 861
OY	241	TGAATGCAAAACAGAAAAATCACCCCATCTTCATCGTTTAGATTGAACAAGTTGCTGACA 300
Dd	862	TGAATGCAAAACAGAAAAATCACCCCATCTTCATCGTTTAGATTGAACAAGTTGCTGACA 921
OY	301	TCACTCCTCAAGAGTTCAGCAAAAAAGTCTCCAGTCCCAAGATGTGTGGACAGCAA 360
Dd	922	TCACTCCTCAAGAGTTCAGCAAAAAAGTCTCCAGTCCCAAGATGTGTGGACAGCAA 981
OY	361	TCAAAATGGCCAACAAGAAAAATGAAGAAGAACATATTCTTGACCATTCACCTGCAT 420
Dd	982	TCAAAATGGCCAACAAGAAAAATGAAGAAGAACATATTCTTGACCATTCACCTGCAT 1041
OY	421	CATGGGATTGGAGAAAAAAGGTGTCTATCACCAAGTAAAGTACCAAGGGGCTGTGAA 480
Dd	1042	CATGGGATTGGAGAAAAAAGGTGTCTATCACCAAGTAAAGTACCAAGGGGCTGTGAA 1101
OY	481	GGGGTGGGGCGTTTTCTGCCACGGGAGCCATTGAAGCAGACATGCAATAGCAACAGAG 540
Dd	1102	GGGGTGGGGCGTTTTCTGCCACGGGAGCCATTGAAGCAGACATGCAATAGCAACAGAG 1161
OY	541	ACCTGTTAGCCCTTCTGAACAAGAATCTGTGACTGTGTGGAAAGAAAGCGAAGTCTTT 600
Dd	1162	ACCTGTTAGCCCTTCTGAACAAGAATCTGTGACTGTGTGGAAAGAAAGCGAAGTCTTT 1221
OY	601	ACAATGATGTCAGATATCAATCGTTTGAATGGGTTTTAGAACATGATGTCATGGCAGCTG 660

[illegible]


```

OY 103 CCAAGTTTACACACAGAAACAGGTGTCTTCACTGTTCCAACTATGAGAGTGAACATG 162
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 38 CCACGTGTGGCAGCCGAGAGAGACCTGATGTCCATGTACGAGAGTGGCTGTGTAACAGC 97
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 163 GACGTGTCAACCTAACACAGAAAGAGCAAGAGACTTGAGATTTCACAAATAACT 222
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 98 GGAAGGTGTACCAACGGCTGGGAGAGAGAGAGAGCGCTTCGATCTTCAGAGCAACCC 157
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 223 CGAATATATCAAGGACATGATGAATCAAAACAGAAATACCCCATCTTCATCGTTAGAT 282
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 158 TGGGATTCATCCAGCAGACCAACTCCGCGAGAGACCGAACTCTCAACG-----TCGGAC 211
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 283 TGAACAAGTTTGGCTGACATCTCTCAAGAGTTTCAGCAAAAAGTACTTGCAGAGTCCCA 342
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 212 TGAACCGTTTGGCTGATCTCACCAACAGAGAAATACAGGGCCAGTACTTGGGAACCAAGA 271
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 343 AGATGTGTGCGCAAAATCAAAATGAGCCACACAGAAATGAAGAAGAACATATTTCTT 402
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 272 TCGATCCCAACGGAGCTGGGAAAGACCCCGCAACCGCTACGGCCACG---TGTG 328
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 403 GTGACCATCCACCTGCATCATGGGATTGGAGAAAAAGGTGTATCACCACAGTAAAGT 462
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 329 GCGACAAATTTGCTGTATCCGTGATTGGAGAGAGAGAGAGTGTCTCTCTGTCAAG 388
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 463 ACCAAGGGGGCTGTGAGAGGGGTTGGCGCTTTCTGCCACGGGAGCCATAGAGACAGCAC 522
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 389 ACCAAGAGAGGCTGTGGAGCTGTGGCATTTTCAGCAATCGGTGACAGAGAAATTA 448
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 523 ATGCATACCAACAGAGACCTTGTTAGCTTTCTGAACACAGACTGCTAGACTGTG--- 579
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 449 ATTAAGTAACTAAACAGGCACTGATTCGTATACAGAACAAAGATGTGGATTTGTGATA 508
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 580 TGAAGAAAGCGAAGGTTTCTTACATGATGGAGATATCAATGCTCGAATGGGCTTTAG 639
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 509 CTGGATATATACCAAGATGCAATGAGAGACTTATGACATATGATTTGATTTATATCA 588
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 640 AACAGTGGGCAATTCCTACTGATGATATTCCTTACAGAGCTAAAGAGGAGTAGATGA 699
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 569 ACAATGGCGCATTTGATCTGATGAGATATACCATACCGTGGTGTGATGGTAGATGGG 628
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 700 AAGCCATTAAGTACAAAGCAAG---GTTACATTTGACGATATGAAGTCTATATATGT 756
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 629 ACACATATAGGAAAAATGCTTAAAGTGTTCATTTGATGACTACGAAAGATGTTCCCTG 688
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 757 CAGATGAGATCAGAAATCAGAGACAGACAGCAAGCTTTTAAAGCGCATCCCTTGAGCAAC 816
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 689 ATGATGAGTT-----AGCCTTGAAGAAAGGCCGTTTGCAGAAATCAGC 727
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 817 CAATTAGTGTCTCAATTGTATGCAAAAGATTTTCAATTTATACACCGGGGGAATTTATGATG 876
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 728 CCGTAGCGCTGCTATTTGAAGGAGGGGCAAGGAATTTCAATTAATG---TATCTGGTG 784
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 877 GAGAAACGTGTACAGTCCGATATGAGATATTCATCTTTGTTTACTTGTGGTATATGTT 936
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 785 TATTCACGGGGAGATGTGCACAGACTAGATCATGCTGTGCGGTGTGGGTATAGAA 844
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 937 CAGCGGATGTGTAGATTTACTGTAGCGAAAAATTCATGGGAGAAAGATTTGGGGAGAG 966
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 845 CAGCTAAAGTGTATGATTTTGTGATGCTTAAAGGAATTCATGGGGTCTTACGTGGGAGAGG 904
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 997 ATGTTTACATTTTGGATCCAAAGAAACACGGGTAA 1030
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 905 ATGGCTACATCAGATTAGAAAGAAATCTGTCTTA 938
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 5
ABN83213
ID ABN83213 standard; cDNA: 1392 BP.

XX ABN83213;
AC
XX
XX 23-SEP-2002 (first entry)
XX

```

DE Soybean D3-beta cDNA.
XX
XX Soybean; angiotensin converting enzyme inhibitor; hypertension;
KM hypotensive; taste; D3-beta; gene; ss.
XX
XX Glycine max.
OS
XX
XX Key
FH 1.1392 Location/Qualifiers
FT CDS
FT FT /tag= a
FT FT /product= "D3-beta"
FT FT /partial
FT FT /note= "No stop codon given"
FT sig_peptide 1..396
FT FT /tag= b
FT mat_peptide 397..1392
FT FT /tag= c
FT FT /note= "Mature D3-beta"
PN
PN W020025546-A1.
PD
PD 18-JUL-2002.
PF
PF 15-JAN-2002; 2002WO-JP00194.
PR
PR 16-JAN-2001; 2001JP-0007400.
PR 04-OCT-2001; 2001JP-0308974.
XX
XX (AJIN ) AJINOMOTO CO INC.
XX
XX Kodera T, Nio N;
XX
XX WPI: 2002-520117/55.
XX P-PSDB: ABB81808.
XX
XX Peptides, useful as hypotensive agents or in health foods -
XX
XX Example 1; Page 30-34; 43pp; Japanese.
XX
XX The invention relates to a novel set of peptides and their salts, which
XX act as inhibitors of angiotensin converting enzyme. The peptides of the
XX CC invention have hypotensive activity. The peptides are used as hypotensive
XX CC agents or in health foods, and have favourable taste. The present
XX CC sequence encodes a soybean D3-beta protein, used in example 1 of the
XX CC invention.
XX
XX Sequence 1392 BP; 364 A; 319 C; 379 G; 330 T; 0 other;
XX
XX
XX Query Match 13.4%; Score 154.4; DB 24; Length 1392;
XX Best local Similarity 52.9%; Pred. No. 3.1e-35;
XX Matches 494; Conservative 0; Mismatches 401; Indels 39; Gaps 6;
OY 103 CCAAGTTTACACAGAAACAGGTGTCTTCACTGTTCCAACTATGAGAGTGAACATG 162
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 104 CCAGTGTGGCAGCCGAGAGAGACCTGATGTCCATGTACGAGCAGTGGCTGTGAACACAG 163
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 163 GACGTGTCAACCTAACACAGAAAGAGCAAGAGACTTGAGATTTCACAAATAACT 222
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 164 GGAAGGTGTACCAACGGCTGGGAGAGAGAGAGAGCGCTTCGATCTTCAGAGCAACCC 223
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 223 CGAATATATCAAGGACATGATGAATCAAAACAGAAATACCCCATCTTCATCGTTAGAT 282
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 224 TGGGATTCATCCAGCAGACCAACTCCGCGAGAGACCGAACTCTCAACG-----TCGGAC 277
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 283 TGAACAAGTTTGGCTGACATCTCTCAAGAGTTTCAGCAAAAAGTACTTGCAGAGTCCCA 342
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 278 TGAACCGTTTGGCTGATCTCACCAACAGAGAAATACAGGGCCAGTACTTGGGAACCAAGA 337
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 343 AGATGTGTGCGCAAAATCAAAATGAGCCACACAGAAATGAAGAAGAACATATTTCTT 402
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 338 TCGATCCCAACGGAGGCTGGGAAAGACCCCGGACAAAGCTTACGGCCACG---TGTG 394
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 403 GTGACCATCCACCTGCATCATGGGATTGGAGAAAAAGGTGTATCACCACAGTAAAGT 462
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```


XX	oil seed rape cysteine protease cDNA clone CDCXS12.
DE	
XX	Promoter; oil seed rape; cDNA clone CDCXS12;
KW	cysteine protease; disrupter protein; plant; tolerance; herbicide;
KM	insect pest; fungal disease; improved yield; improved quality;
KW	fertility control; ss.
OS	
XX	Brassica napus.
XX	
FH	Key
FT	location/Qualifiers
CDS	30..1361
FT	/transl_except= (pos:690..692, aa:Xaa)
FT	/note= "Xaa = unknown"
FT	/*tag= a
XX	
PN	WO9735983-A2.
XX	
PD	02-OCT-1997.
PF	
PE	18-MAR-1997; 97WO-GB00729.
XX	
PR	22-MAR-1996; 96GB-0006062.
XX	
PA	(ZENE) ZENECA LTD.
XX	
PI	Greenland AJ, Jepson I, Thomas DRP;
XX	
DR	WPI: 1997-489646/45.
DR	P-PsDB: AAM27441.
XX	
PT	Cysteine protease promoter - obtained from oil seed rape, useful to develop plants with improved agronomic characteristics
XX	
PS	Claim 18; Fig 12; 137pp; English.
XX	
CC	The present sequence encodes an oil seed rape cysteine protease,
CC	the promoter for which can restrict the expression of a disrupter
CC	protein gene to a suitable stage of plant development, to provide
CC	plants with novel agronomic features, e.g. tolerance to herbicides,
CC	insect pests and fungal diseases, improved yield and/or quality of
CC	harvested product, and novel mechanisms for the control of plant
CC	fertility.
XX	
SO	Sequence 1577 BP; 458 A; 325 C; 356 G; 437 T; 1 other;
	Query Match 13.1%; Score 151.2; DB 18; Length 1577;
	Best Local Similarity 55.2%; Pred. No. 3e-34;
	Matches 369; Conservative 1; Mismatches 274; Indels 24; Gaps
OY	414 CCTGCATCATGGGATTGGAGGAAGAAAGGTGTCATCACCACCAAGTAAGTACCAAGGGCGC 473
DB	402 CCAGATTCTGTTGATTGGAGGAAGAAGAGGTGTACTAATGTCCAAGATCAAGGAAGC 461
OY	474 TGTCGAAGGGTGTTGGCGCGTTTCTCGCACGGAGAGCCATGAGACACAGATGCATAGCA 533
DB	462 TCGCGAGCGGTGTTGGTCTTCTCGGGAGCATGTGAGCTATGAAGAGATCAACCAAGATTGTA 521
OY	534 ACAGAGACCTTGTTAGACCTTTCTGACAAGAATCTGTAGACTGTGTGGAAGAACGGA 593
DB	522 ACAGAGGATCTCATCAGCCTCTCTGSCAAGAACATCATGATTGTATGAATCATATCAAC 581
OY	594 GGTTCTTCAATGAGATGGC--AGTATCAATCGTTGATGGTTTTAGAACATGGTGGC 650
DB	582 GATGATGCAATGCGTGGTCTCATGACATRCGCTTTTCAATTTGTCTATTAAAACCATGGG 641
OY	651 ATTGCCACTGATGATGATTATCCTTACACAGCGTAAAGAGGGTAGATGCAAAGCCAATAAG 710
DB	642 ATTTGACACAGAAAGATATATCCTTATCAAGAACGATGGGACCGCTGMAAGAAAGATTA 701
OY	711 ATACGAAGACAGGTTTCAATTGACGATATGAACCTCTAATATGTCAAGTGAAGTACA 770
DB	702 TTGAATTAGAAAGGTGTGCAATTTGATGCTACCGCTGATGAAGAAATCAATGACGAGAA 761

Accession	Gene	Protein	Length	Source			
Oy	771	GAATCAGACAGACAGCAGAGCGTTCTTAAGGGCCATCCCTTGACCAACCAATTAAGTGTCA	830	Arabidopsis thaliana			
Db	762	GCCTTACTAGAAAGCTGTAGCGGCTCAGCCAGTTAGTGTG-----GTATTCCT	809	Arabidopsis thaliana			
Oy	831	ATTGATCCAAAAGATTTTTCATTTATATACACCGGCAATTTATGATGAGAAACATGTACA	890	Arabidopsis thaliana			
Db	810	GGGAGCGAGAGAGCGCTTACGTTACTTATCTCTAAGGGAAATTTCTCTGGCCATGTTCAACA	869	Arabidopsis thaliana			
Oy	891	AGTCGGTATGGGATTAATCACTTTGTTTACTTGTGGTTATGTTTCAGCGGATGGTGA	950	Arabidopsis thaliana			
Db	870	TCATTTG-----GACCACGACAGTGCATCTGTGTGGATTAGGGTTCAAGAACGTTGT	920	Arabidopsis thaliana			
Oy	951	GATTACTCGATAGAGAAATTTTCATTTGGGGAGAAAGATTGGGAGAAAGATGTTACATTTGG	1010	Arabidopsis thaliana			
Db	921	GATTACTCGATCGTGAAGAACTTTGGGGAAAGAGTTGGGGAAATGATGATGGTTATCCAC	980	Arabidopsis thaliana			
Oy	1011	ATCCAAAGAAACACAGCGGTAAATTTATTTAGAGAGTGTGGGATGAATTTATTTTCCTTCATAC	1070	Arabidopsis thaliana			
Db	981	ATGCAAGGTTACACACCGGCACGCGAAGAGATATGCGGAATCAACATGCTGCTTCATAT	1040	Arabidopsis thaliana			
Oy	1071	CCAACCA 1078		Arabidopsis thaliana			
Db	1041	CCCATCAA 1048		Arabidopsis thaliana			
RESULT 9							
AT90170	standard; cdna; 1553 bp.						
Id	AT90170	standard; cdna; 1553 bp.					
xx	AA90170;						
xx	20-APR-1998	(first entry)					
DE	Oil seed rape cysteine protease cdna clone CDCYS14.						
xx							
KW	Promoter; oil seed rape; cDNA clone CDCYS14;						
KW	cysteine protease; disrupter protein; plant; tolerance; herbicide;						
KW	insect pest; fungal disease; improved yield; improved quality;						
KW	fertility control; ss.						
OS	Brassica napus.						
xx							
FH	Key	Location/Qualifiers					
FT	CDS	41..1369					
FT	/*tag= a						
xx							
PN	W09735983-A2.						
xx							
PD	02-OCT-1997.						
xx							
PE	18-MAR-1997; 97MO-GB00729.						
xx							
PR	22-MAR-1996; 96GB-0006062.						
xx							
PA	(ZENEC) ZENECA LTD.						
xx							
PI	Greenland AJ, Jepson I, Thomas DRP;						
xx							
DR	WPI: 1997-489646/45.						
DR	P-PSDB; AAW27442.						
xx							
PT	Cysteine protease promoter - obtained from oil seed rape, useful to						
PT	develop plants with improved agronomic characteristics						
xx							
PS	Claim 19; Fig 13; 137pp; English.						
xx							
CC	The present sequence encodes an oil seed rape cysteine protease,						
CC	the promoter for which can restrict the expression of a disrupter						
CC	protein gene to a suitable stage of plant development, to provide						
CC	plants with novel agronomic features, e.g. tolerance to herbicides,						
CC	insect pests and fungal diseases, improved yield and/or quality of						
CC	harvested product, and novel mechanisms for the control of plant						

PN	ED1033405-72.	
PD	06-SEP-2000.	
PF	25-FEB-2000;	2000EP-0301439
XX	25-FEB-1999;	99US-0121825
PR	05-MAR-1999;	99US-0122380
PR	09-MAR-1999;	99US-0125348
PR	23-MAR-1999;	99US-0125788
PR	25-MAR-1999;	99US-0126264
PR	20-MAR-1999;	99US-0126785
PR	01-APR-1999;	99US-0127462
PR	06-APR-1999;	99US-0128234
PR	08-APR-1999;	99US-0128714
PR	16-APR-1999;	99US-0129645
PR	19-APR-1999;	99US-0130077
PR	21-APR-1999;	99US-0130449
PR	23-APR-1999;	99US-0130510
PR	23-APR-1999;	99US-0130891
PR	28-APR-1999;	99US-0131449
PR	30-APR-1999;	99US-0132048
PR	30-APR-1999;	99US-0132407
PR	04-MAY-1999;	99US-0132484
PR	05-MAY-1999;	99US-0132485
PR	06-MAY-1999;	99US-0132486
PR	06-MAY-1999;	99US-0132487
PR	07-MAY-1999;	99US-0132683
PR	11-MAY-1999;	99US-0132456
PR	14-MAY-1999;	99US-0134219
PR	14-MAY-1999;	99US-0134318
PR	14-MAY-1999;	99US-0134321
PR	14-MAY-1999;	99US-0134327
PR	18-MAY-1999;	99US-0134370
PR	19-MAY-1999;	99US-0134941
PR	20-MAY-1999;	99US-0135124
PR	21-MAY-1999;	99US-0135353
PR	24-MAY-1999;	99US-0135629
PR	27-MAY-1999;	99US-0136021
PR	28-MAY-1999;	99US-0136392
PR	10-JUN-1999;	99US-0136782
PR	10-JUN-1999;	99US-0138540
PR	10-JUN-1999;	99US-0138847
PR	14-JUN-1999;	99US-0139119
PR	16-JUN-1999;	99US-0139452
PR	16-JUN-1999;	99US-0139453
PR	17-JUN-1999;	99US-0139457
PR	18-JUN-1999;	99US-0139454
PR	18-JUN-1999;	99US-0139455
PR	18-JUN-1999;	99US-0139456
PR	18-JUN-1999;	99US-0139461
PR	18-JUN-1999;	99US-0139463
PR	18-JUN-1999;	99US-0139750
PR	18-JUN-1999;	99US-0139763
PR	21-JUN-1999;	99US-0139617
PR	22-JUN-1999;	99US-0139699
PR	23-JUN-1999;	99US-0140053
PR	23-JUN-1999;	99US-0140354
PR	24-JUN-1999;	99US-0140625
PR	28-JUN-1999;	99US-0140823
PR	29-JUN-1999;	99US-0140991
PR	30-JUN-1999;	99US-0141847
PR	01-JUL-1999;	99US-0141847

PR 01-JUL-1999; 99US-0142154.
 PR 02-JUL-1999; 99US-0142055.
 PR 06-JUL-1999; 99US-0142390.
 PR 08-JUL-1999; 99US-0142803.
 PR 09-JUL-1999; 99US-0142820.
 PR 12-JUL-1999; 99US-0142977.
 PR 13-JUL-1999; 99US-0143542.
 PR 14-JUL-1999; 99US-0143624.
 PR 15-JUL-1999; 99US-0144005.
 PR 16-JUL-1999; 99US-0144085.
 PR 16-JUL-1999; 99US-0144086.
 PR 19-JUL-1999; 99US-0144325.
 PR 19-JUL-1999; 99US-0144331.
 PR 19-JUL-1999; 99US-0144332.
 PR 19-JUL-1999; 99US-0144333.
 PR 19-JUL-1999; 99US-0144334.
 PR 19-JUL-1999; 99US-0144335.
 PR 20-JUL-1999; 99US-0144352.
 PR 20-JUL-1999; 99US-0144632.
 PR 20-JUL-1999; 99US-0144684.
 PR 21-JUL-1999; 99US-0144814.
 PR 21-JUL-1999; 99US-0145086.
 PR 21-JUL-1999; 99US-0145088.
 PR 22-JUL-1999; 99US-0145085.
 PR 22-JUL-1999; 99US-0145087.
 PR 22-JUL-1999; 99US-0145192.
 PR 23-JUL-1999; 99US-0145145.
 PR 23-JUL-1999; 99US-0145218.
 PR 23-JUL-1999; 99US-0145224.
 PR 26-JUL-1999; 99US-0145276.
 PR 27-JUL-1999; 99US-0145913.
 PR 27-JUL-1999; 99US-0145918.
 PR 27-JUL-1999; 99US-0145919.
 PR 28-JUL-1999; 99US-0145951.
 PR 02-AUG-1999; 99US-0146386.
 PR 02-AUG-1999; 99US-0146388.
 PR 02-AUG-1999; 99US-0146389.
 PR 03-AUG-1999; 99US-0147038.
 PR 04-AUG-1999; 99US-0147204.
 PR 04-AUG-1999; 99US-0147302.
 PR 05-AUG-1999; 99US-0147192.
 PR 05-AUG-1999; 99US-0147260.
 PR 06-AUG-1999; 99US-0147303.
 PR 06-AUG-1999; 99US-0147416.
 PR 09-AUG-1999; 99US-0147493.
 PR 09-AUG-1999; 99US-0147935.
 PR 10-AUG-1999; 99US-0148171.
 PR 11-AUG-1999; 99US-0148319.
 PR 12-AUG-1999; 99US-0148341.
 PR 13-AUG-1999; 99US-0148365.
 PR 13-AUG-1999; 99US-0148684.
 PR 16-AUG-1999; 99US-0149368.
 PR 17-AUG-1999; 99US-0149175.
 PR 18-AUG-1999; 99US-0149426.
 PR 20-AUG-1999; 99US-0149722.
 PR 20-AUG-1999; 99US-0149723.
 PR 20-AUG-1999; 99US-0149929.
 PR 23-AUG-1999; 99US-0149902.
 PR 23-AUG-1999; 99US-0149930.
 PR 25-AUG-1999; 99US-0150566.
 PR 26-AUG-1999; 99US-0150884.
 PR 27-AUG-1999; 99US-0151065.
 PR 27-AUG-1999; 99US-0151066.
 PR 27-AUG-1999; 99US-0151080.
 PR 30-AUG-1999; 99US-0151303.
 PR 31-AUG-1999; 99US-0151438.
 PR 01-SEP-1999; 99US-0151930.
 PR 07-SEP-1999; 99US-0152363.
 PR 10-SEP-1999; 99US-0153070.
 PR 13-SEP-1999; 99US-0153758.
 PR 15-SEP-1999; 99US-0154018.
 PR 16-SEP-1999; 99US-0154039.

PR 20-SEP-1999; 99US-0154779.
 PR 22-SEP-1999; 99US-0155139.
 PR 23-SEP-1999; 99US-0155486.
 PR 24-SEP-1999; 99US-0155659.
 PR 28-SEP-1999; 99US-0156458.
 PR 29-SEP-1999; 99US-0156596.
 PR 04-OCT-1999; 99US-0157117.
 PR 05-OCT-1999; 99US-0157753.
 PR 06-OCT-1999; 99US-0157865.
 PR 07-OCT-1999; 99US-0158029.
 PR 08-OCT-1999; 99US-0158232.
 PR 12-OCT-1999; 99US-0158369.
 PR 13-OCT-1999; 99US-0159293.
 PR 13-OCT-1999; 99US-0159294.
 PR 13-OCT-1999; 99US-0159295.
 PR 14-OCT-1999; 99US-0159329.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 14-OCT-1999; 99US-0159638.
 PR 18-OCT-1999; 99US-0159584.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160981.
 PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161992.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query Match 11.8%; Score 136.6; DB 21; Length 1083;
 Best Local Similarity 50.4%; Pred. No. 5, 3e-30;
 Matches 405; Conservative 0; Mismatches 374; Indels 24; Gaps 2;

QY 283 TGAACAAGTTGCTGACATCTCTCAAGAGTTTCAGCAAAAGTAC-----TTGCCAAG 336
 DB 296 TAAATCAGTTGCTGATTTAAACCAATGACGAAATTTGTCACATGTAAGTGTTCACAAAG 355
 QY 337 CTCCCAAGGATGTGCGACCAATCAAAATGCGCAACAAGAAAGTGCATCACCACAT 396
 DB 356 GTGTCTCGCATTTATCTAGCCCAAGCCCAACTAAATGTCGCCGTTTAGTACCAAAAGC 415
 QY 397 ATTCTGTGACCATCCACCTGCATCATGGATGAGAAAAAGGTGCATCACCACAT 456
 DB 416 TTCTCTGCTGCTTTCGCGGTTTCTGTGACTGAGAAAGAGAGCTGTGACCCCTA 475
 QY 457 TAAAGTACCAAGGGGCTGTGGAAGGGGTTGGCGCTTTCTGACAGGAGCCATATGAAG 516
 DB 476 TCAAGATCAAGGACCTGCGGATGTTGTTGGCGTTTTCAGCGGTTTGGCGATTGAAG 535
 QY 517 CAGCACAATGATAGCAACAGAGACCTTGTAGCCTTCTGAAACAAGAACTGTAGACT 576
 DB 536 GAGCAACACAATAATGAAGAAAGGAACTTATATCTTTGTCAGACACAGCTTTGTTGATT 595
 QY 577 GTGTGAGAGAAAGCAAGGTTTACATATGATGCGCATATCAATGTTTCGATGGGTTT 636
 DB 596 GCGACACAAAGCATTTTGGCTGGAAGGGGCTTATATGATACGCTTTAGACATATAA 655
 QY 637 TAGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 696
 DB 656 AAGCGACTGGCGCTTGCACACAGTCAATTAATCTTACAAAGGCAAGCGTACTT 715

PA (FLET-) FLETCHER CHALLENGE FORESTS LTD.
XX
PI Film B, Lasham A;
XX
DR WPI: 2001-061724/07.
DR P-PSDB; AAB85766.
XX
PT Novel defender against cell death polynucleotide useful for modulating
PT programmed cell death pathway and specific development pathways in
PT forestry plant -
XX
PS Claim 1: Page 86; 142pp; English.
XX
CC The present invention relates to coding sequences (see AAF44740-F44840
CC and AAF44843-F44844) and proteins (see AAB65714-B65814) involved in
CC programmed cell death (PCD; apoptosis). The coding sequences and proteins
CC of the present invention are useful for modulating a PCD or cell death
CC pathway and various developmental pathways in a forestry plant, by
CC stably incorporating one of the present coding sequences into the genome
CC of the forestry plant, where the coding sequence provides a PCD pathway
CC that is not present in a native form of the forestry plant.
XX
SQ Sequence 1600 BP; 390 A; 420 C; 446 G; 344 T; 0 other;

Query Match 11.5%; Score 132.8; DB 22; Length 1600;
Best Local Similarity 51.0%; Pred. No. 8.7e-29;
Matches 466; Conservative 0; Mismatches 412; Indels 36; Gaps 5;

OY 124 AGGTCTTCTACTGTTCCAACTATGCAAGATGACATGACGTGTCTACCATRACCAAG 183
DB 231 AGGTGATGCGCTCTACGAGAGCTGGCTGCCAAGCGGCAAGGCTTACAGCCCTGG 290
OY 184 AAGAGAGGCAAGAGACTGAGATTTCAGAAATATCTGAATATATCAGGAGCATGA 243
DB 291 GCGAGAGAGAGAGCGCTTCAGAGTCTTCAAGACACCTCCGGTTCATGACACACCA 350
OY 244 ATGCAACAGCAAAATACACCCATCTCATCTGTTAGATTGAACAATGTTGTCACATCA 303
DB 351 ACGCGCGCGGGGA-----CCGAGCTACACGGTGGCGCTCAACCAATGTCGCGACCTCA 404
OY 304 CTCTCTCAAGGTTTACGAAAAAGTACTTGCACAGCTCCCAAGATGTGTCCAGCAATCA 363
DB 405 CTAACGAGGAGTACCGGTTCATGATCTGAGCGCCAGAGATGATCGCGGCGCGCGC 464
OY 364 AATGCGCAACAGAAATATGAAGAGCAATATTTCTTGACATCCACCTGCATCAT 423
DB 465 TCGGAGAGGCGCCGACGATCGGTAGCCCTGGCGCGGAGAGACTGCGCGCTCG 524
OY 424 GGAATTGAGAGAAAAAGGTGTATCACCCAGTAAGTAAAGTACCAAGGGGCTGTGAAAGG 483
DB 525 TCATTTGAGAGGAAGAGGCGCGCTGTGATGACGTCAAGGACCAAGGAGGAGGAGTT 584
OY 484 GTTGGGCGTTTGTGCCACGCGGACCATAGACGACATGCAATATGCAACAGAGAAC 543
DB 585 GGTGGGCGTTCTTACCAATGTCGTCTGAGAGGATTAACAAGCTGTGACGCGTATT 644
OY 544 TTGTTAGCCCTTCTGAACAAGACTGTAGATGTGGAAG---AAAGCAAGGTTCTT 600
DB 645 TGATCTCTCTGTCGACGACAGAACTGTGACCTGCGATATCTCTCAATAGGAAGATGA 704
OY 601 ACAATGAGTGGCAATATCAATCGTGCATGGTATTGAACATGTGTGCATGCCACTG 660
DB 705 ATGGCGGCGCTCATGATTAATGCTTTGATTTATCAACAGGAGGATTTGATCGG 764
OY 661 ATGATGATTTCTCTTACAGAGCTAAAGAGGGTGAAGT---CAAAAGCAATATGATCAAG 717
DB 765 AGGAAGATTATCCCTATAGAGCTGTAGATGCACTGTGACCAATACAGAGAACGCAA 824
OY 718 ACAAGGTTAATGATGAGATGAACCTATTAATATGATGATGATGATGATGATGATG 777
DB 825 AGGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 872
OY 778 AGACAGAGCAAGCGTTCTTAAGCGCCATCTTGAGCAACCAATTAAGTCTCAATTGATG 837

DB 873 -----AAGCATGCAAAAGGAGGAGTTCCTAATCAACAGTACAGTGGCCATTGAAG 923
OY 838 CAAAAGATTTTCAATTATATACCCGGGGAATTTATGATGAGAAACTGTACAGTCCGT 897
DB 924 CAGAGGCGCGGAATTCACATTTTATG---ATTGCGTATATTTACTGCAAAATGCGCA 980
OY 898 ATGAGATTAATCACTTGTGTTTACTTGTGCTTATGATGATGATGATGATGATGATG 957
DB 981 CAGCTCTGATCATGCGGTTACTGACAGTGTGATATGCGACAGAAACGAGTTGATTACT 1040
OY 958 GGAATGCAAAATTCATGCGGAGAGAGTGGGAGAGATGGGAGACATGCTTACATTTGATCCAA 1017
DB 1041 GCATAGTGAAGAACTCATGCGCGGTAGTGTGCGAGAGCAAGGCTTACATCAAAATGGCAC 1100
OY 1018 GAAACAGCGGTAT 1031
DB 1101 GAAATGTGCCAAT 1114
RESULT 14
AAC40001
ID AAC40001 standard; DNA: 1698 BP.
AC
XX AAC40001;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 26686.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EF1033405-A2.
XX
PD 06-SEP-2000.
XX
PE 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.

PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140921.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142820.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.

PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149302.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 25-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-015138.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 11, 2003, 19:13:12 : Search time 2663 Seconds
(without alignments)
12633.437 Million cell updates/sec

Title: US-09-805-694B-1

Perfect score: 1156

Sequence: 1 gcgcgcgcgtggttccctt.....ctctcttgcgcgcgcgcgc 1156

Scoring table:

IDENTITY_NUC

Gapop 10.0 , Gapept 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:*
1: gb_ba:*
2: gb_hgt:*
3: gb_in:*
4: gb_cm:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_scs:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_cm:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_scs:*
28: em_un:*
29: em_vl:*
30: em_hgt_hum:*
31: em_hgt_inv:*
32: em_hgt_other:*
33: em_hgt_mus:*
34: em_hgt_pln:*
35: em_hgt_rtd:*
36: em_hgt_mam:*
37: em_hgt_vrt:*
38: em_sy:*
39: em_hgtgo_hum:*
40: em_hgtgo_mus:*
41: em_hgtgo_other:*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1156	100.0	1156	6	AX252299	AX252299 Sequence
2	1156	100.0	2970	6	AX252300	AX252300 Sequence
3	1140	98.6	1287	8	SORP34A	J05560 Glycine max
4	475.4	41.1	5449	8	AB013290	AB013290 Glycine m
5	469.6	40.6	6675	8	AB013289	AB013289 Glycine m
6	251.2	21.7	1732	6	AF182079	AF182079 Matricari
7	167	14.4	1807	8	E13051	E13051 cDNA encodi
8	160.6	13.9	1755	8	ST0245924	AJ245924 Solanum t
9	159.4	13.8	1756	8	LEA3137	AJ003137 Lycopersi
10	159.4	13.8	1779	8	AF172856	AF172856 Lycopersi
11	155.4	13.4	1751	8	AF259983	AF259983 Ipomoea b
12	155.2	13.4	1699	8	AF454957	AF454957 Brassica
13	154.4	13.4	1056	6	AR049578	AR049578 Sequence
14	154.4	13.4	1056	6	E13053	E13053 Modified CD
15	154.4	13.4	1056	6	E14274	E14274 DNA encodin
16	154.4	13.4	1849	6	E13052	E13052 cDNA encodi
17	151.2	13.1	1577	6	A65336	A65336 Sequence 59
18	151.2	13.1	1577	6	AR150474	AR150474 Sequence
19	150.6	13.0	1680	8	PVZ999954	Z99954 Phaseolus v
20	149.6	12.9	1361	8	VSCYSPRO	X75749 V.sativa L.
21	147	12.7	1553	6	A65337	A65337 Sequence 60
22	147	12.7	1553	6	AR150475	AR150475 Sequence
23	146	12.6	1638	8	AF388175	AF388175 Arabidops
24	144	12.5	1485	8	AY114661	AY114661 Arabidops
25	144	12.5	1628	8	AY062608	AY062608 Arabidops
26	142.2	12.3	1284	8	VSCYSPROA	Z34895 V.sativa L.
27	142	12.3	1389	8	AY081730	AY081730 Arabidops
28	142	12.3	1420	8	AY133781	AY133781 Arabidops
29	142	12.3	1604	8	AY039606	AY039606 Arabidops
30	142	12.3	1656	8	AY072130	AY072130 Arabidops
31	141.4	12.2	1351	8	PVZ999952	Z99952 Phaseolus v
32	141.4	12.2	1382	8	PVA224766	AJ224766 Phaseolus
33	141.4	12.2	1382	8	PVU52970	U52970 Phaseolus v
34	139.2	12.0	1219	8	AF191028	AF191028 Arabidops
35	138	11.9	1041	8	AY040073	AY040073 Arabidops
36	137	11.9	1218	8	NRCYSPRO	Z99173 Nicotiana t
37	136.6	11.8	1253	8	AF370131	AF370131 Arabidops
38	135.4	11.7	1414	8	AY093350	AY093350 Arabidops
39	135.4	11.7	1575	8	AY062725	AY062725 Arabidops
40	135	11.7	1425	8	AY043294	AY043294 Arabidops
41	133.6	11.6	1314	8	VMSHEP	X15732 Vigna mungo
42	133.2	11.5	1830	8	AF454958	AF454958 Brassica
43	132	11.4	1221	8	PVPEP1	X56753 Phaseolus v
44	131.8	11.4	1820	8	PSYH10L	X66061 P.sativum m
45	130.8	11.3	1221	8	VR049445	U49445 Vigna radia

ALIGNMENTS

RESULT 1
AX252299
LOCUS AX252299 1156 bp
DEFINITION Sequence 1 from Patent WO0168887.
ACCESSION AX252299
VERSION AX252299.1 GI:15985640
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 1156)
AUTHORS Jung, R. and Kinney, A.J.
TITLE Hypoallergenic transgenic soybeans
JOURNAL Patent: WO 0168887-A 1 20-SEP-2001;
E. I. du Pont de Nemours and Company (US) ; PIONEER HI-BRED

QY 361 TCAGAAATGGCCACACAGAAAAATGAGAGAACAAATATCTTGTGACCATCACCCTGCAT 420
 |||||||
 Db 982 TCAGAAATGGCCACACAGAAAAATGAGAGAACAAATATCTTGTGACCATCACCCTGCAT 1041
 |||||||
 QY 421 CATGGGATTTGGAGAGAAAAAGGTGTCTACCCCAAGTAAAGTACCAAGGGGCGTGTGAA 480
 |||||||
 Db 1042 CATGGGATTTGGAGAGAAAAAGGTGTCTACCCCAAGTAAAGTACCAAGGGGCGTGTGAA 1101
 |||||||
 QY 481 GGGGTTGGGCGTTTTCGCCACGGAGCCATAGAGACGACATGCATATGCAACACAGAG 540
 |||||||
 Db 1102 GGGGTTGGGCGTTTTCGCCACGGAGCCATAGAGACGACATGCATATGCAACACAGAG 1161
 |||||||
 QY 541 ACCCTGTAGCCCTTCTGAAACAAGAACTCTAGACTGTGTGGAAGAAAGCGAAGTTCTT 600
 |||||||
 Db 1162 ACCCTGTAGCCCTTCTGAAACAAGAACTCTAGACTGTGTGGAAGAAAGCGAAGTTCTT 1221
 |||||||
 QY 601 ACAATGGATGGCAGTATCAATCGTTCGAATGGGTTTGAACATGTGTGCGATTGCCACTG 660
 |||||||
 Db 1222 ACAATGGATGGCAGTATCAATCGTTCGAATGGGTTTGAACATGTGTGCGATTGCCACTG 1281
 |||||||
 QY 661 ATATGATATATCTTACAGAGCTAAAGAGGCTAGATGCAAGCCAAATAGATACAGACA 720
 |||||||
 Db 1282 ATATGATATATCTTACAGAGCTAAAGAGGCTAGATGCAAGCCAAATAGATACAGACA 1341
 |||||||
 QY 721 AGCTTACAAATTTGACGAGTATGAAACTCTAAATATGTCAGATGAGAGTACAGAAATCAGAGA 780
 |||||||
 Db 1342 AGCTTACAAATTTGACGAGTATGAAACTCTAAATATGTCAGATGAGAGTACAGAAATCAGAGA 1401
 |||||||
 QY 781 CAAGAGCAAGCGTTCTTAAGCGCCATCCTTACAGACCAATTAAGTGTCTCAATTTGATGCA 840
 |||||||
 Db 1402 CAAGAGCAAGCGTTCTTAAGCGCCATCCTTACAGACCAATTAAGTGTCTCAATTTGATGCA 1461
 |||||||
 QY 841 AAGATTTTCAATTTATACACGGGGGAATTTATGATGAGAGAAACTGTACAAAGTCCGATG 900
 |||||||
 Db 1462 AAGATTTTCAATTTATACACGGGGGAATTTATGATGAGAGAAACTGTACAAAGTCCGATG 1521
 |||||||
 QY 901 GGATTAATCACTTTGTTTACTTGTGGTTATGTTTACAGCGGATGTGTAGATTACTGGA 960
 |||||||
 Db 1522 GGATTAATCACTTTGTTTACTTGTGGTTATGTTTACAGCGGATGTGTAGATTACTGGA 1581
 |||||||
 QY 961 TACCGAAAAATTCATGGGGGAGAGATTGGGAGAAAGATGTGTACATTTGATCCAAAGAA 1020
 |||||||
 Db 1582 TACCGAAAAATTCATGGGGGAGAGATTGGGAGAAAGATGTGTACATTTGATCCAAAGAA 1641
 |||||||
 QY 1021 ACACGGGTAATTTATAGAGTGTGTGGATGATTTTTCGGTCTACCAACCAAG 1080
 |||||||
 Db 1642 ACACGGGTAATTTATAGAGTGTGTGGATGATTTTTCGGTCTACCAACCAAG 1701
 |||||||
 QY 1081 AGGAATCAGAAACACTGTGTCTGCTCGCTTAAAGGTCATCGAAGAGTTGATCCTCTC 1140
 |||||||
 Db 1702 AGGAATCAGAAACACTGTGTCTGCTCGCTTAAAGGTCATCGAAGAGTTGATCCTCTC 1761
 |||||||
 QY 1141 CTCTTTGAGCGGCC 1156
 |||||||
 Db 1762 CTCTTTGAGCGGCC 1777
 |||||||

RESULT 3
 SORP34A
 LOCUS 1287 bp mRNA linear PLN 30-SEP-1996
 DEFINITION glycine max 34 kDa maturing seed vacuolar thiol protease mRNA,
 complete cds.
 ACCESSION J05560
 VERSION 105560.1 GI:1199562
 KEYWORDS soybean protein storage vacuole thiol protease.
 SOURCE glycine max.
 ORGANISM Glycine max.
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
 Glycine.
 1 (bases 1 to 1287)
 REFERENCE
 AUTHORS Kalinski,A., Weisemann,J.M., Matthews,B.F. and Herman,E.M.

TITLE Molecular cloning of a protein associated with soybean seed oil
 JOURNAL J Biol. Chem. 265 (23), 13843-13848 (1990)
 MEDLINE 90338001
 PUBMED 2380191
 REFERENCE 2 (bases 1 to 1287)
 AUTHORS Kalinski,A., Melroy,D.L., Dwivedi,R.S. and Herman,E.M.
 TITLE A soybean vacuolar protein (P34) related to thiol proteases is
 synthesized as a glycoprotein precursor during seed maturation
 J Biol. Chem. 267 (17), 12068-12076 (1992)
 JOURNAL MEDLINE 92291086
 PUBMED 1601876
 REFERENCE 3 (bases 1 to 1287)
 AUTHORS Herman,E.M.
 TITLE Direct Submission
 JOURNAL Submitted (18-MAY-1990) Eliot M. Herman, Plant Molecular Biology
 Laboratory, USDA, Agricultural Research Service, Beltsville, MD
 20705
 COMMENT On Feb 21, 1996 this sequence version replaced g1:170039.
 FEATURES
 source Location/Qualifiers
 1..1287
 /organism="Glycine max"
 /strain="Century"
 /db_xref="taxon:3847"
 /clone="p34/1350"
 /tissue="seed"
 <1..1287
 3..1142
 /codon_start=1
 /product="34 kDa maturing seed vacuolar thiol protease
 precursor"
 /protein_id="FAB09252.1"
 /db_xref="GI:1199563"
 /translation="MGFVLLFLSLGSSSSISTHRSILDLDTLTKPTTQKVSLE
 QLMKSEHGRVYHNHEEFAKRLIEFKNSNTIRMANNRKSPSHRGLKFPADITPQE
 FSKRYLQAPRDVSOQIKMANKRKKEQYSCDHPASMDNRKGVITQVYQGGCGRW
 AFSATGAIENAHAIATGDLVSLSEQDVESEPSYMGOMQSEFVYHGGIATD
 DQYPRKRGRCRANKIODKVTIDGVEITLMSDESEFEEOAFSAIIIEOPISVID
 AKDFHLYTGGIYDNCSTPFGCIHNFVLLVYGSAAGVDVWIKNSGEMDKEDEYIW
 IQRNTGNLGLVCQMNFFASTPKRESEETIVSAKVRKRDHNSPL"
 3..368
 sig.peptide
 mat.peptide
 369..1139
 /product="34 kDa maturing seed vacuolar thiol protease"
 BASE COUNT 407 a 241 c 297 g 342 t
 ORIGIN

Query Match 98.6%; Score 1140; DB 8; Length 1287;
 Best Local Similarity 100.0%; Pred. No. 7.2e-286;
 Matches 1140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 ATGGGTTCCCTTGTTGCTCTTTTCTCCCTTACGTCCTCTTCTAGTTCACGATA 68
 |||||||
 Db 3 ATGGGTTCCCTTGTTGCTCTTTTCTCCCTTACGTCCTCTTCTAGTTCACGATA 62
 |||||||
 QY 69 TCAACTCATGCTTCATATTGAGACCTTGACCTTAACCAAGTTTACACAGAAACAGGTG 128
 |||||||
 Db 63 TCAACTCATGCTTCATATTGAGACCTTGACCTTAACCAAGTTTACACAGAAACAGGTG 122
 |||||||
 QY 129 TCTTCACTGTTCCAACTATGAGAGATGAGCATGAGCTGTCTACCAACAGAGAA 188
 |||||||
 Db 123 TCTTCACTGTTCCAACTATGAGAGATGAGCATGAGCTGTCTACCAACAGAGAA 182
 |||||||
 QY 189 GAGGCAAGAGACCTTGAGATTTCACAAATATCTCGAATATATCAGGAGACATGATGCA 248
 |||||||
 Db 183 GAGGCAAGAGACCTTGAGATTTCACAAATATCTCGAATATATCAGGAGACATGATGCA 242
 |||||||
 QY 249 AACGAAATACACCCCTTCATCGTTTAGATTGAAAGTTGCTGCATCATCTCT 308
 |||||||
 Db 243 AACGAAATACACCCCTTCATCGTTTAGATTGAAAGTTGCTGCATCATCTCT 302
 |||||||
 QY 309 CAAGAGTTCAAGCAAAAGATCTTGCAAGCTCCCAAGAGATGTGTGCGACGAATCAAAATG 368
 |||||||
 Db 303 CAAGAGTTCAAGCAAAAGATCTTGCAAGCTCCCAAGAGATGTGTGCGACGAATCAAAATG 362
 |||||||

LOCUS	DEFINITION	AB013290	5449 bp	DNA	Linear	PLN 29-APR-1998
AB013290	Glycine max pseudogene for Bd 30K.					
ACCESSION	AB013290					
VERSION	AB013290.1					
KEYWORDS	pseudogene; Bd 30K.					
SOURCE	Glycine max DNA.					
ORGANISM	Glycine max					
REFERENCE	Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.					
AUTHORS	1 (bases 1 to 5449) Takano,T. and Yamada,M.					
TITLE	Cloning of a pseudo gene for an allergenic protein in Gly m Bd 30K in soybean					
JOURNAL	Published Only in Database (1998)					
REFERENCE	2 (bases 1 to 5449)					

AUTHORS	Takano, T. and Yamada, M.
TITLE	Direct Submission
JOURNAL	Submitted (22-Apr-1998) Tetsuo Takano, University of Tokyo, Asian Natural Environmental Science Center; University Farm, Univ. Tokyo 1-1-1 Midori-cho, Tanashi, Tokyo 188-0002, Japan (E-mail:takanoe@n.u-tokyo.ac.jp, Tel:+81-424-63-1618, Fax:+81-424-63-1618)
FEATURES	
source	Location/Qualifiers 1..5449 /organism="Glycine max" /db_xref="taxon:3847"
misc-feature	1..5449 /note="Pseudogene for an allergenic protein Gly m Bd 30k"
BASE COUNT	1823 a 917 c 862 g 1847 t
ORIGIN	
Query Match	41.1%; Score 475.4; DB: Length 5449;
Best Local Similarity	71.7%; Pred No. 6.4e-113;
Matches	802; Conservative 0; Mismatches 101; Indels 216; Gaps 5
OY	9 ATGGGTTCCTTGTTGCCTTCTTCCCTTAGAGTCCTCCTTAATGAACGACATA 68
Dd	3421 ATGGGTTCCTTGTTGCCTTCTTCCCHCTTAGAGTCCTCCTTAATGAACGACATA 3480
OY	69 TCAACTCATCGTTCCATATTGACCTTGACCCTAACCAAGTTTACCAACAAGAAGGTG 128
Dd	3481 T---CTCATCGTTCCATATTGACCTTGACCCTAACCAAGTTTACCAACAAGAAGGTG 3537
OY	129 TCCTTACGTGCCAAGTATGAAAGGTGACATGAGCAGCTGTACCATTAACACGAAGA 188
Dd	3538 TCCTTACGTTCCTCAACCATGAAAGGTAGACACGACGAGGCTTACCATTAACAGAGA 3597
OY	189 GAGGCAAGAGACTTGAGATTTTCAAAGATAAAGTCCAAGTATATCACGAGCATGAATGA 248
Dd	3598 TAGGCAAGAGACTTGAGATTTTCAAAGATAAAGTCCAAGTATATCACGAGCATGAATGA 3657
OY	249 AACCAAAAATCCCCCATTCGATCGTTAGAGTTGAACAAAGTTTGTCACATCCCTCT 308
Dd	3658 AACCAAAAATCCCCCATTCGATCGTTGAGATTGAACAAAGTTTGTCACATCCCTCT 3717
OY	309 CAAGAGTTCAGCAAAAAAGTACTTGCAAGCTCCCACAGAGTGTGCGACGCAAAATG 368
Dd	3718 CAAAGTTCAGCAAAAAAGTATTGCAAGCTCCCACAGAGTGTGCGACGCAAAATG 3777
OY	369 GCCAACAGAAAATGAAGACGAACAATATTTTGTGACATCCACCTGCATCATGGAT 428
Dd	3778 GCCAACAGGAATTTAAGAGGAGAACAGCATTCCTTGTGACATCCACCTGCATCCTGGGAT 3837
OY	429 TGAGAGAAAAAGGTTGTATCCACCAAGTAAGTACCAAGGGGGGCT----- 476
Dd	3838 TGAGAGAAAAAGGTTGTATCCACGAGTGAAGCACCAAGGCTTGTGTGACGTGACATG 3897
OY	477 ----- 476
Dd	3898 AAACCATGTTTTAGTACGTAAATTAATTAACACTAGTATTTTCTCTCTTT 3957
OY	477 -----GGAAGGGGTGGGGCTTTCTGCCA 501
Dd	3958 GA AAAATGACTTGTGAGTGTGA AACCATT TTGATAGGGA GTGGGATTTGTGTCGA 4017
OY	502 CGGGAGCATATAGACAGACATGCAATATGCAACAGAGACCTGTTAGCTTTGAGAC 561
Dd	4018 CGGGAGCATATAGACAGCTGATCGAATATGCAACAGAGACCTTG--TACCCTTTTGAGAC 4076
OY	562 AAGAACTCGTAGACTGTGTGGAAGAAGCGAAGTCTTTCACATGATGATCAAT 621
Dd	4077 AAGAACTCGTAGATGTGTG--GAAGGAGGATGTGTTCCAGTGGAGAGCCCGATGATT 4133
OY	622 CGTTGGAATGGTTTAAACAATGGTGGCATTTGCCACGTATATGATTAATCTTACAGAG 681
Dd	4134 CGTTGGAATGGTTTAAACAATGGTGGGATTTGCCACGTATATGATTAATCTTACAGAG 4193

QY 682 CTAAGAGGGTAGATGCAAAAGCCAATA----- 708
|||||
Db 4194 CTAAGAGGTGTAGATGCAAAAGCCAATAAGTCTGTGAATATATAGCGAATATGCA 4253
QY 709 ----- 708
Db 4254 TGCATTCCTTTTGAATTAATTAAACCTTCATCTCTTTTGTGCACTTTTTTTCCT 4213
QY 709 -----AGATTAAGACCAAGGTATACAAATTGACGGATATGAAGCTTAATATGTCAG 759
|||||
Db 4314 TCTGGAGTAGATGATACAGACAGAGTTTCAATTTACGATATGCAATATATAGATGCCAT 4373
QY 760 ATGAGAGTACAGATGAGAGACAGACAGACCGTCTTAAAGCCCATCTTGTAGCAACCA 819
4374 ATGAGAGTACAGATGAGAGAGAGAGAACCGTCTTATGCGCCATCTTGTAGCAACCA 4433
QY 820 TTGAGTCTCAATGATGCAAAAGATTTCATTTATACACCGGGGAATTTATGATGAG 879
4434 TTAGTGTCTCCATTTGATGCAAGAGATTTCATCTCACATGGGGTAATAATATATG 4493
QY 880 AAAAGTACAGTCCGTATGAGTAAATCATCTTTGTT 918
Db 4494 TACTGTCTTCAAAAAATAAGTGTATCTTACATTTTAT 4532

RESULT 5
AB013289 6675 bp DNA linear PLN 29-APR-1998
LOCUS AB013289
DEFINITION Glycine max gene for Bd 30k, complete cds.
ACCESSION AB013289
VERSION AB013289.1 GI:3097320
KEYWORDS Bd 30k.
SOURCE Glycine max DNA.
ORGANISM Glycine max

REFERENCE 1 (bases 1 to 6675)
AUTHORS Takano,T. and Yamada,M.
TITLE Cloning of a gene for an allergenic protein in soybean
JOURNAL Published Only in Database (1998)
REFERENCE 2 (bases 1 to 6675)
AUTHORS Takano,T. and Yamada,M.
TITLE Direct Submission
JOURNAL Submitted (22-APR-1998) Tetsuo Takano, University of Tokyo, Asian
Natural Environmental Science Center; University Farm, Univ. Tokyo
1-1-1 Midori-cho, Tanashi, Tokyo 188-0002, Japan
(E-mail:takano@ims.u-tokyo.ac.jp, Tel:+81-424-63-1618,
Fax:+81-424-63-1618)
FEATURES
source location/Qualifiers
1. 6675
/organism="Glycine max"
/db_xref="taxon:3847"
join(4647..5115,5474..5706,5820..5972,6178..6462)
/codon_start=1
/product="Bd 30k"
/protein_id="BAA25899.1"
/db_xref="GI:3097321"

BASE COUNT 2349 a 1024 c 1074 g 2228 t
ORIGIN
Query Match 40.6%; Score 469.6; DB 8; Length 6675;
Best Local Similarity 99.2%; Pred. No. 2e-111;
Matches 472; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 9 ATGGGTTTCCTTGTGTCCTTTCTTTCTCCCTTAGAGTCTCTTCTAGTCCACATTA 68
|||||
Db 4647 ATGGGTTTCCTTGTGTCCTTTCTTTCTCCCTTAGAGTCTCTTCTAGTCCACATTA 4706
QY 69 TCAACTCATCTGTTCCATTTGACCTTGGACCTTACCAAGTTTACCAACAGAAACAGGTG 128
4707 TCAACTCATCTGTTCCATTTGACCTTGGACCTTACCAAGTTTACCAACAGAAACAGGTG 4766
QY 129 TCTTCATCTGTTCCACATTTGAGAGAGTGGACATGAGCGTGTCTACATTAACCCGAGAA 188
4767 TCTTCATCTGTTCCACATTTGAGAGAGTGGACATGAGCGTGTCTACATTAACCCGAGAA 4826
QY 189 GAGGCAAGAGACTTGAATTTTCAAGAAATTAACCTCAATATATACAGGACATGAATGA 248
4827 GAGGCAAGAGACTTGAATTTTCAAGAAATTAACCTCAATATATACAGGACATGAATGA 4886
QY 249 AACAGAAATCACCACATTCATCTGTTTAAAGATGAACAAAGTTTCTGACATCACTCT 308
4887 AACAGAAATCACCACATTCATCTGTTTAAAGATGAACAAAGTTTCTGACATCACTCTCT 4946
QY 309 CAAGAGTTCGCAAAAAGTACTGCAAGCTCCCAAGAGTGTGCGAGCAAAATCAAAATG 368
4947 CAAGAGTTCGCAAAAAGTACTGCAAGCTCCCAAGAGTGTGCGAGCAAAATCAAAATG 5006
QY 369 GCCACAGAAATGAAGAAGACAAATATCTTGTGACCATCACCCTGCATCATGGAT 428
5007 GCCACAGAAATGAAGAAGACAAATATCTTGTGACCATCACCCTGCATCATGGAT 5066
QY 429 TGGAGGAAAAAGGTGTATCACCACCAAGTAACTACCAAGGGGCTGTGAGAGGG 484
5067 TGGAGGAAAAAGGTGTATCACCACCAAGTAACTACCAAGGGGCTGTGATGTG 5122

RESULT 6
AF182079 1732 bp mRNA linear PLN 10-SEP-1999
LOCUS AF182079
DEFINITION Matricaria chamomilla thiol protease (ctp) mRNA, complete cds.
ACCESSION AF182079
VERSION AF182079.1 GI:5853328
KEYWORDS
SOURCE Matricaria chamomilla.
ORGANISM Matricaria chamomilla

REFERENCE 1 (bases 1 to 1732)
AUTHORS Kohchi,C., Yasuda,M. and Hirata,T.
TITLE Isolation of cDNA encoding thiol protease from cultured shoot
JOURNAL unpublished
REFERENCE 2 (bases 1 to 1732)
AUTHORS Kohchi,C., Yasuda,M. and Hirata,T.
TITLE Direct Submission
JOURNAL Submitted (01-SEP-1999) Department of Mathematical and Life
Sciences, Graduate School of Science, Hiroshima University, 1-3-1
Kagamiyama, Higashi-Hiroshima 739-8526, Japan
FEATURES
source location/Qualifiers
1. 1732
/organism="Matricaria chamomilla"
/db_xref="taxon:98504"
/tissue_type="cultured shoot primordia"
1. 1732
/gene="ctp"
44..1549
/gene="ctp"
/codon_start=1
/product="thiol protease"
/protein_id="AAD54424.1"
/db_xref="GI:5853329"

BASE COUNT 1024 a 1074 c 1074 g 2228 t
ORIGIN
Query Match 40.6%; Score 469.6; DB 8; Length 6675;
Best Local Similarity 99.2%; Pred. No. 2e-111;
Matches 472; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OQGGCSCMAFVSGSGISASNAIATGDTIATSECLLVLCDDPYDADSCGGNDTATATWII
 KNGLSDSEDDVPIYTSNSGRDCKDRTSAKVVLSLVEYEVESNEDCALVATTPATTII
 IGIYGAISAFOLYLTGYGYNGCCSSKPYDIDHVAIVYGSGSDGDWYIKNSMSWTYI
 LEGYIIMERNMDIKNGVCGMYLEPIYVPTTAPRPGPPPPAPPPPPPPPPAPPA
 SKCDPFYKCADORTCCCIPEFYNYCLJYSGKSYSDAVCCNKSAAACSCSDYICDVGNG
 YCYNSAKTFEVPARKROLAKHKMAKMEIETETIKEDRPLAMNENPFMAAA"

Query Match	21.7%	Score	251.2	DB	8	Length	1732
Best Local Similarity	53.4%	Pred. No.	1.6e-54				
Matches 572; Conservative	0	Mismatches	448			Indels	12
						Gaps	4

OY	70	CAACCTACACGTCCTCATCTTGGACCTTGACCTAACCAAGTTTCCACACAGCAAGACGGTGT	129
Db	123	CTAGTAGATTTTCTATCTCTTGAAGGTCACAGAAATGATATTTCTATTAAGTGCACAAAGTTA	182
OY	130	CTTCACTGTTCCAACTATGGAGAAGTAGAGCATGAGCGTGTACACATAACCCAGAAAG	189
Db	183	GTGACCTATTTGGAAATATGGAAAGAGTGTGCATGGAAAAACATACCAACTGAAAGAAAG	242
OY	190	AGCGAAGAGACTTGAGATTTTCAAGAAATPACTCGAAGTATATCGAGGACATGAATGCA	249
Db	243	AGAACCTTAGCGTTGAGAAATTTCAAGAAAAGTGAAGCTTTGTAATGGGAAAAAAGCTG	302
OY	250	ACAGAAATACCCCACTTCTCATCGTTTGGATTTGAATGAACAGTTTGCTGCATCACTCTC	309
Db	303	AGAGAAAGTCTGAGCTAGATCATACGTGTGGAGTGAATAAGTTTGTCTAGCTGAATATG	362
OY	310	AAGAGTTACGCAAAAAGTACTTGCAGAGCTGCCAAG--ATGTGTGCGACGCAATCAAAA	366
Db	363	AAGAGTTTAAAGAGTATATATGTCAAAGTTAAAGGGGCTGTGAAGTATGATTTAAAGA	422
OY	367	TGCGCAACAGAAATATGAAGAGCAACATATTTCTTG--TGACCATCCACCTGCATAT	423
Db	423	TGGGGGGGTGTAAAAAGGAACATAGTAGTAAAGTTGAGAGCACTGTGTATGCCAACCTTT	482
OY	424	GGGATTTGAGAAAAAAGGTGTCATCACCCAAAGTAAAGTACCAAGGGGGCTGTGGAAGG	483
Db	483	TGGATTTGGAGAGCAAAAGAGTAGTTACACATGAAAGGCCAAGGCCAGTGTGGAATTT	542
OY	484	GTTGGGGTTTTCTGCCAGGGAGCCATGAAGACACACATGCATAGCAACAGGACCC	543
Db	543	GTTGGGCAATTTCTCAGTGTCTGGATCATGTGAAGAGTGCAAATGCATATAGCCAGGTGACC	602
OY	544	TTGTAGACCTTTTGAACAAGAACTCGTAGACTGTGTGAAGAAAAGCGAAGGTTCTTACA	603
Db	603	TCATTTAGACTCTCAGACCAAGACCTTGTGCAGTGTACACTATGACTATAGGTGCCATG	662
OY	604	ATGCAATGGCAATATCAATCGTTGCAATGGGTTTTAGACATGTGTGCATTTGCCACTGATG	663
Db	663	GTGGAACAATGAGACAGCGCTTATAGATGAATTAATTAAGATATGGTGGCTGTATCTTGAG	722
OY	664	ATGATTTATTCCTTACAGAGCTTAAGAGGGTATGATGCAAAACCATTAAGATTAACAAGCAAG	723
Db	723	ACGATTTCCATACACCAAGTTCTTAATGGCG----CGATGGTAAATGTGACAAACAAAG	778
OY	724	TTACAATTTAGCGATATGAACCTTAATAATGTCAAGATGAGAGTACAGAAATCAGACAG	783
Db	779	TCACCAAGAAGTCAGTTGCTGCTGTGATACCTATGTGGAAGTGAATTCGAATGAATGACA	838
OY	784	AGCAAGGTTCTTAAGGGCCATCCTTGAGCAACCAATTAGTGTCTCAATTTGATGCAAAG	843
Db	839	GTTTTATGTGTGTGGCACTACCCCTGTTA--CTATTGGTATCTGTGGTGTGGCTTTATG	896
OY	844	ATTTTCTATTATACACCGGGGGAATTTATGATGAGAAACATGTACAAATGCCGTGTGCA	903
Db	897	ATTTTCCAGTATACACCGGGGGGTATATTAAGAGCAATAGCTCAAGTAAACCATATGATA	956
OY	904	TTAATCACTTTGTTTACTTGTGGGTTATGTTCACGCGAGTGTGTAGATTACTGATAG	963
Db	957	TAGACCAATGACGTCTATATGCTGTGTATGTTCACAAAGATGGCAAGGACACTATGATGC	1016

QY	964	CGAAAAATTCTATGGGAGAAAGTTGGGGACAGATGTTCATCTTTGGATCCAAAGAACA	1023
Dd	1017	TCAAGAAATTCATGGGCCACTTATTTGGGAGCATGAAGGCATCATTCTAATGGAAAAGAAC	1076
QY	1024	CGGGTAATTTTATGAGAGTGTTGGGATGATATTTCCCTTCATACCACCAACAAAGG	1083
Dd	1077	CTGATATTTAAAAACGGGGTTTGGAATGTACCTTGAGCGGGTGTACCCTATCACTGCGG	1136
QY	1084	AATCAGAAACAC	1095
Dd	1137	CGCCAACACCC	1148

LOCUS	1807 bp	DNA	linear	PAT 27-APR-1998
EL3051				
DEFINITION	cDNA encoding thiol protease D3-alpha.			
ACCESSION	EL3051			
VERSION	EL3051.1	GI:3251863		
KEYWORDS	JP 1997121870-A/1.			
SOURCE	Glycine max.			
ORGANISM	Glycine max. Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.			
REFERENCE	1 (bases 1 to 1807)			
AUTHORS	Kawai,M., Asano,M., Suzuki,S., Miwa,T., Shibai,H. and Yokozeki,K.			
TITLE	DNA CAPABLE OF CODING NEW THIOL PROTEASE AND PRODUCTION OF THE SAME			
JOURNAL	THIOL PROTEASE USING THE DNA Patent: JP 1997121870-A 1 13-MAY-1997; AJINOMOTO CO INC			
COMMENT	OS Glycine max PN JP 1997121870-A/1 PD 13-MAY-1997 PF 28-DEC-1995 JP 1995353931 PR 29-DEC-1994 JP 94P 340399, 30-AUG-1995 JP 95P 245279 PI KAWAI MISAKO, ASANO MINAO, SUZUKI SHUNICHI, MIWA TETSUYA, PI SHIBAI HIROSHIRO, YOKOZEKI KENZO PC C12N15/09,C12N1/19,C12N1/21,C12N9/50,(C12N1/19,C12R1:865), PC (C12N1/21, PC C12R1:19),(C12N9/50,C12R1:19),(C12N9/50,C12R1:865); CC strandedness: Double; CC Topology: linear; CC hypothetical: No; CC anti-sense: No; FH Key Location/Qualifiers FT source 1..1807 FT /organism='Glycine max' FT 5'UTR 1..67 FT sig-peptide 68..460 FT /product='signal peptide of thiol protease FT FT D3-alpha' 461..1456 FT mat-peptide /product='thiol protease D3-alpha' FT FT 68..1459 /product='thiol protease D3-alpha' FT 3'UTR FT 1460..1807. FEATURES source Location/Qualifiers 1..1807 /organism='Glycine max' /db_xref='taxon:3847' BASE COUNT 500 a 387 c 449 g 471 t ORIGIN Query Match 14.4% Score 167; DB 6; Length 1807; Best Local Similarity 53.0%; Pred. No. 1.3e-32; Matches 493; Conservative 0; Mismatches 405; Indels 33; Gaps 5; 103 CCAAGTTACCAACAGAAACAGGTGCTCTCACTTCCAACTATGCAAGAGTGAGCATG 162			

D	b	485	GTTGGGCAATTCCTCCTCCCGTGTGCATAGAAATCAATAAATCCGATTAGTAACCGGAAT	544
O	y	544	TTGTTTAGCCCTTCTCTAACAAGAAGACTCGTAGACTG---TGSAAAAAGCCAGGTTCTT	600
D	b	545	TGATATCATCTATCACAGAGAGAGTGGTGAGATGATGAAGTCGTACATAGTAAGGTTGCG	604
O	y	601	ACAATGGATGGCAGATCAATGTTCTGAAATGGGCTTTTGAACATATGSGGCANTGGCACMG	660
D	b	605	ATGTGTGCTTATGAGACTACGCCCTTTGAAATTCGTATCATATATGAGAGAAATGTGCACCTG	664
O	y	661	ATGATGATTATTCCTTACAGAGAGCTAAAGAGGGTATGCAAAACCATAATGATACAGACA	720
D	b	665	AAGAAGACTAACCTTACAAAAGAACGTATATGACGTATGGGACCAATATAGGAAAAATGGCA	724
O	y	721	AGGTTACAAATTGACGGATATGAACCTCTAATATATGTCAAGATGAGAGTACAGATACAGAA	780
D	b	725	AGGTTGTTAA-----AATGATAGCATATGMAAATGTTCCGTTAATA	766
O	y	781	CAGACCAAGCGTTCTTAAGCCGCATCCTTGAGCAACCAATATGTCATATGATATGAA	840
D	b	767	ACGAAAGGCGGTACAAAAGGCTGTTGCACATCAACCTGTGAGCAATGCAATGGAAGCG	826
O	y	841	AAGATTTTCATTTATACACCGGGGGAATTTATGATGAGAGAAACGTACAATCCGATG	900
D	b	827	GTGGCAGAGACTTAACAGCAC---TACAATCTGATATCTTACTGGAANAATGSGTACTG	883
O	y	901	GGATTAATACACTTTGTTTACTTGTGGGTTATGTTGACAGCGGATGCTAGATTACTGGA	960
D	b	884	CAGTGCATCATGCTGTTACTTGTCTGCTGATATGCTTGAGATGCGATGATGATTATTGGA	943
O	y	961	TAGCAAAAATTCATAGGGGAGAGAGTTGGGGAGAGACATGTTACATTTTGATCCAAAGA	1020
D	b	944	TTGTTAGGAACATCAGGGAGGCTAATATGGGAGAGAAAGGCTACCTCAGAGTCCAACGA	1003
O	y	1021	ACACGGTATATTTATAGGAGTGTGGGATCAATATTTGCTCTCATACCCACCAAGA	1080
D	b	1004	ACGTTGCCACGCTCTAGTGGCTTGTGTGGTTTAGCCACAGAGCCTTCATATCCAGTAAGA	1063
O	y	1081	AGGAATCAGAAACACTGCTGCTGCTC	1107
D	b	1064	CAGGCGCAAATCCTCTTAACCCGCTC	1090
R	E	S	RESULT 9	
L	E	A	LEAJ3137	
L	O	C	LOCUS	
D	E	F	DEFINITION	
A	C	C	Accession	
K	E	Y	Version	
W	O	R	Keywords	
S	O	U	Source	
O	R	G	Organism	
R	E	F	Reference	
A	U	T	Title	
J	O	U	Journal	
R	E	F	Reference	
A	U	T	Authors	
T	I	T	Title	
J	O	U	Journal	
F	E	A	Features	
S	O	U	Source	
			location/Qualifiers	
			1..1756	
			/organism="Lycopersicon esculentum"	
			/cultivar="Cherry"	
			/db_xref="taxon:4081"	

gene	29..1429
CDS	/gene="C14" 29..1429 /gene="C14" /function="cysteine protease" /codon_start=1 /product="Cyp1" /protein_id="CA05894.1" /db_xref="GI:2828252" /db_xref="SP:EMBL:049877" /translation="MAHSSSTLPISILMLIFSTLSASDMSIISYDETHIHRTDE VSALEFEMLIHEGKSYNALGEKDRPFIIPDNRIYIDEQSNVQSYKGLTFRPADLT NEEARSYLIGTSSGDRKLTSKNSSDYLVKVDSDLSIPEDIMREKGYLVKQSGSG SCAERSVVAAMESINAIIVTNLSISQELVDDRSINBECCDGLMVAEPYIKNG IDBEDPIKERNSVCDQYIRKNAKYVAIDTEDYPAENKALDRAVAHDSYALEBG GRFOHQKSGILFTEGCGTAADHGVLAGYETENGMDWIYIRNSGAWMGVLTLYOR NVASSGLIGLATEPSYPVKTPGNPFPAPSPSPVYPAPCEDEYSQCAVGTCCCL OFKRSCEGMCPLGATCCEDHDHSCCPHDYPICNVROGCSMSKGNPLKMKRIL AQPICFNGNGKKSS"
BASE COUNT	540 a 344 c 406 g 466 t
ORIGIN	
Query Match	13.8%; Score 159.4; DB 8; Length 1756;
Best Local Similarity	50.8%; Pred. No. 1,2e-30;
Matches	506; Conservative 0; Mismatches 461; Indels 30; Gaps 4;
Db	114 ACACAGAAACAGGTGCTCTTCACCTGTTCCAACTATGGAAGTGAAGCATGAGCTGTCTAC 173 149 ACCGACGATGAAGTCTCGCGCTTGACGAGTCGCGTGAATCGACAGCAAGAAATCGTAC 208
Qy	174 CATACCAAGAAAGAGAGCAAAAGACTTGATGATTTCAAGATTAACCTGCACTATATC 233 209 AACGCTTAGAGGAAAGAAAGATGAAGATTTCAGATCTTCAAGATTAACCTGATGATAC 268
Qy	234 AGGAGCATGATGCAAAACAGAAATCACCCCATCTCATCGTTAGATGAAACAAGTTT 293 269 GAGCAACAGAACTCTGTTCCGAT-----CAGATTTTAAGTCGCGATTACCGAAATTC 322
Db	294 GCTGACATCACTCTCAAGAGTTGAGCAAAAAAGTACTTGAAGCTCCCAAGATGTGTG 353 323 GCCGATGTGACTAATGAGAGAGTACGTCGATTTACTTGGGAACGAAGAAGTTCCGGTGAC 382
Qy	354 CAGCAATCAAAATGGCCACACAGAAATGAAGAAGCAACTATTTCTTGACCATCCA 413 383 CGGAAAGATTAATCAAGAACAAAGATATCGGATTCCTCTAAAGTTGGGGAATAGCTTG 442
Qy	414 CCTGCATATGAGATTTGAGAGAAAAAAGGTGTCATCACCCAAAGTAAAGTACCAAGGGGC 473 443 CCGGATCAATTTGACTGAGAGAAAAAAGGTGTGCTTTGTTGTCACAGATCAAGGAAGC 502
Db	474 TGTGGAAGGGGTGGGCTTTCTGCGCAAGGGAGCCATAGAAGCAGCATGACATAGCA 533 503 TGTGGAAGTTGTTGGGCAATCTCTGCTGTGCTGATGATCAATTAACGCGATATTC 562
Qy	534 ACAGAGACCTTTGAGCTTTCTGTAACAAGAACTCCGTAGACTGTG---TGGAGAAAGC 590 563 ACTGGATTTGATATCTACTATCAAGACGAAGAAGTTGGATTGTGATAGGTCGTACAT 622
Qy	591 GAAGGTTCTTCAATGATGAGGAGATCAATCGTTGCAATGGGTTTGGATCACTGGTGC 650 623 GAGGTTCCGATGGTGTCTTATGGACTACGCCCTTTGATTTCTGTCATTTAACAATGGAGGA 682
Db	651 ATTGCACCTGATGATGATATCTCTTACAGAGCTAAAGAGGGGTAGATGCAAGCAATPAG 710 683 ATCGACACTGAAGAAGCTACCCCTTACAAACAAGCAATGGCGTATGTGATCATATPAGG 742
Qy	711 ATCAAGACAAAGTTACAATTGAGGAGATGAAACTCTAATTAATGTGATGAGAGTACA 770 743 AAAATGCCAAGGTGTGTA-----AATAGATAGCTTGAAGATGTT 784
Db	771 GAATCAGACAGAGCAAGCGTTCTTAAGGCCCATCTTGAGCAACCAATTAATGTCTCA 830 785 CCGTGAATTAACGAAGAAGCGCTTGCAGAAAGGCTGTGACATCAACCTGTAGGATTTGA 844

QY	831	ATTGATGCAAAAAGATTTTCATTTATACACCGGGGAATTTATGATGAGAAAACGTAC	890
Db	845	CTTGACAGCTGGTGGCAGACACTCCAGAC--TCAATCTGATATCTTCACTCGAAAA	901
OY	881	AGTCCGATATGGGANTTAATCACTTTGTTTTACTGTGGGTATATGTTACCGCAGATGTGA	950
Db	902	TGTGTCTGCAAGTCATGATCATGTGTACTATTATGCTGATATGATCATGACAAATGGCATG	961
OY	951	GATPACTGATAGCGCAAAAATTCATGGGGAGAAAGTTGGGGAGAGATGTTACTTTGG	1010
Db	962	GATTATTTGGATTCGTTAGCAACTCATGGGAGCTAATCTGGGGAGAAAGCGCTACTCAGA	1021
OY	1011	ATCCAAAGAACACCGGTATATTTATTAGAGTGTGTGGATGATGAATTTATTTGCTTCATAC	1070
Db	1022	GTCAGAGCTTAAGTTGCCAGCTCATGTGCGTTGTGTGTTTAAAGCATAGACGCTTCATAT	1081
OY	1071	CCAACCAAGAGGAATCAGAAACACTGCTGTCTC	1107
Db	1082	CCAGTAAACACGACCAAAATCTCTTAACCCGCTC	1118
RESULT 10	AF172856	1779 bp mRNA linear	PLN 17-AUG-2001
LOCUS	AF172856	Lycopersicon esculentum	cysteine protease TDI-65 (tdi-65) mRNA,
DEFINITION	complete cds.		
ACCESSION	AF172856		
KEYWORDS	AF172856.1	GI:5726640	
SOURCE	Lycopersicon esculentum.		
ORGANISM	Lycopersicon esculentum		
REFERENCE	1 (bases 1 to 1779)		
AUTHORS	Harrak,H., Azelmat,S., Baker,E.N. and Tabaeizadeh,Z.		
TITLE	Isolation and characterization of a gene encoding a drought-induced		
JOURNAL	Genome 44 (3), 368-374 (2001)		
MEDLINE	21337597		
PUBMED	11444695		
REFERENCE	2 (bases 1 to 1779)		
AUTHORS	Harrak,H. and Tabaeizadeh,Z.		
TITLE	Direct Submission		
JOURNAL	Submitted (28-JUL-1999) Biological Sciences, University of Quebec		
	in Montreal, P.O. Box 8888 Station 'Centre Ville', Montreal, Quebec		
	H3C 3P8, Canada		
FEATURES	location/Qualifiers		
source	1..1779		
gene	/organism="Lycopersicon esculentum"		
	/db_xref="taxon:4081"		
	1..1779		
	/gene="tdi-65"		
	48..1448		
	/gene="tdi-65"		
	/note="induced by drought; localized in the nuclei and		
	chloroplast (Tabaeizadeh, Z. et al., 1995. Protoplasma,		
	186:208-219)"		
	/codon_start=1		
	/product="cysteine protease TDI-65"		
	/protein_id="AAD48496.1"		
	/db_xref="GI:5726641"		
	/translation="MAHSHSLTITSLILMIFSTLSSASDKSITISYEPHIHRTTDEE		
	VSALYESWLIETHGKSYNALGEKDRFQIFKDNLAYIDEQNSVPQSYKLGITRKADLT		
	NEEYRSYLTGKSSGDRKKLSKNSKSDYLPKVGSDPESIDMREKGLVGVKDKGSCG		
	SCAFPSVAAMESINAIYVTGNLISLSQELVDCDRSYNEGCGGLMAYAEFVYKNGG		
	IDFEDYPKERNGVCDQYRKNARVVRIDSEDEVVNNKALOKAVAHOPYSTALEAG		
	GRDPEYKSGIFTEGCGTAVDGHGVIAVGYCTGENMDQVWYIYRNSGAWGNGYLROR		
	NVMSSGICGLAIPRSYPVVTGPRPRPARSPSPYKPRPIECDEYSOCAYGTTCCIL		
	QFRSGSGWCCPLEGATCCEDHISCCPHDPIICNVROGTCSSMKGNPLGVKAKRKIL		
	AQPLGCAWNGGKKSSS"		

BASE COUNT	543 a	356 c	410 g	470 t
ORIGIN				
Query Match	13.8%;	Score 159.4;	DB 8;	Length 1779;
Best Local Similarity	50.8%;	Pred. No. 1.2e-30;		
Matches 506;	Conservative	0;	Mismatches 461;	Indels 30;
			Gaps	4.
OY	114	ACACGAAACAGGTGCTTCACTGTTTCCAACTATGGAAGAGAGATGGACGTAGCTGTCTAC	173	
DB	168	ACCGACGATGAAGTCTCGCGCTTTGACAGTCTCTGGCTATCGACGACGGAATAATCGTAC	227	
OY	174	CATAACCAAGAAAGAGGCAAGAGACTTGGAGATTTCTAAGAATATACGAACTATATC	233	
DB	228	AACGCTTAAGAGAAAGAGATTAAGAGATTTCAAGTCTTCAAGATATCTGAGATACATA	287	
OY	234	AGGCACTGAATGCAAAACAGAAATCACCCCATTTCTCATCGTTTAGAATTGAACAAGTTT	293	
DB	288	GACGACAGAACTCTGTTCCGAAT-----CAGAGTTATACCTGGATTGACGAATAATTC	341	
OY	294	GCTGACATTCACCTCCCAAGAGATTCAGCAAAAAGTACTTGCAGCTCCCAAGATGTGTG	353	
DB	342	GCCGATCTGACTAAATAGAGAGATACAGTCTGATTTTACTTTGGGAACGAAGATTCCGGTGCAC	401	
OY	354	CAGCAATCAAAATATGCCCAAGAAAGAAATGAGAAGAGAACATATTTCTGTGACCATCA	413	
DB	402	CGGAAAATGTTATCGAAGACAAAAGTAGATCGGTATCTTCTTAAGTGGGGATACCTTG	461	
OY	414	CTGTCACTGAGGATTTGGCAAAAAAGGTGTCATCCACCAAGTAAGATACCAAGGGGCG	473	
DB	462	CCGGAATCAATTTGACTGAGAGAAAAAGGTGTGCTTTGGTGTCAAGATCAGAGAAC	521	
OY	474	TGTGGAAGGGTGTGGCGCTTCTTGCCACGGAGCCATAGAAAGCAGACATGCATATGA	533	
DB	522	TGTGGAGATTTTGTGGGCATTTCTCTGTTGCTGCCATGGAATCAATTAACGGGATATGC	581	
OY	534	ACAGGAGACCTTGTTAGCCTTCTGAACAAGAACTGCTGACTGTG---TGAGAAAGAC	590	
DB	582	ACTGGGAATTTATATCACTATACAGGACAGAGTTGTGTGATGTGATATGTCGTACAT	641	
OY	591	GAAAGTTCTTACAAATGCATGCGCAGTATCAATGTTTGAATGGGTTTATGAACATGGTGC	650	
DB	642	GAAAGTTGGAGTGGTGTCTTATGGACATCGCCTTTGAATGTGCTATTAAAGATGGAGA	701	
OY	651	ATTGCCACTGATGATGATTTATCTCTTACAGAGCTTAAAGAGGATGATGCAAAAGCAATAG	710	
DB	702	ATCGACACTGAAGAGACTATCCCTTTACAAGAACGCAATGGCGTATGATCAATATAGG	761	
OY	711	ATACAGAACAGGTATACAAATTGACGATTTGAACCTTAAATATGTCACATGAGATACA	770	
DB	762	AAAAATGCCCAAGTTGTTTNA-----AATGATATGCTATGAAAGATGTT	803	
OY	771	GAATCAGAGACAGACCAAGCGTTCTTAAAGCGCATCTTGACACCAATTAATGTGCTCA	830	
DB	804	CCTGTTAATTAACGAAAGAGCGTTTGCAAAAGCGTGTGGCAATCAACCTGTGACATTTGA	863	
OY	831	ATTGATGCAAAAGATTTTCAATTATACACCGGGGGAATTTATGATGAGAAAACTGTACA	890	
DB	864	CTTGAGAGCTGGTGGAGAGACATTTCCAGCAC--TACAATATCGGTATCTTCACTGGAAAA	920	
OY	891	AGTCCGTATGGATTAATCACTTTGTTTACTTGTGGGTTATGTGTTCAACGGATGTGTGA	950	
DB	921	TGTGTACTGTGAGTATCATGTGTAGTATTTATGTGGGATATGATCTGACATGTGCATG	980	
OY	951	GATTTCAGTATGACAAAATTCATGGGGAGAAAGATTTGGGGAGAAAGATGTTACATTTGG	1010	
DB	981	GATTATTGATCTTAAAGACTCAATGGGGAGGCTTAATCTGGGGAGAGAACGGCTTACCTCA	1040	
OY	1011	ATCCAAAGAAACACGCGTAAATTATTTAGAGAGTGTGGGATCAATATATTTCCCTTCATAC	1070	
DB	1041	GTCCAGCGTAAAGCTTCCAGCTCTAGTGGCTTGTGAGTTTACCATTAAGGCTTCATATAT	1100	
OY	1071	CCAACCAAGAGGATTCAGAAACACTGGTGTCTGCTC	1107	

DB	1101	CGAGTAAACAGCAGCAAAATCCTCTCTTAACCCCGCTC	1137
RESULT	11		
LOCUS	AF255983	1751 bp	mRNA linear
DEFINITION	Ipomoea batatas putative cysteine protease mRNA, complete cds.		
ACCESSION	AF255983		
VERSION	AF255983.1	GI:13897889	
KEYWORDS			
SOURCE			
ORGANISM	Ipomoea batatas. Euphorbia Euphorbia		
REFERENCE	1 (bases 1 to 1751) Chen, H.-J., Hou, W.-C. and Lin, Y.-H.		
AUTHORS	Isolation of a cDNA clone encoding a putative cysteine protease from senescent leaves of sweet potato (Ipomoea batatas cv. Tainong 57)		
TITLE	Unpublished		
JOURNAL	2 (bases 1 to 1751)		
REFERENCE	Chen, H.-J., Hou, W.-C. and Lin, Y.-H.		
AUTHORS	Direct Submission		
TITLE	Submitted (24-APR-2000) Academia Sinica, Institute of Botany, Nankang, Taipei, Taiwan 11529, Republic of China		
JOURNAL	Location/Qualifiers		
FEATURES	1..1751		
Source	/organism="Ipomoea batatas"		
	/cultivar="Tainong 57"		
	/db_xref="taxon:4120"		
	/tissue_type="senescent leaf"		
	28..1416		
	/codon_start=1		
	/product="putative cysteine protease"		
	/protein_id="AAK48495.1"		
	/db_xref="gi:13897890"		
	/translation="MAKLLILSLFVLAAVSASASADMSITTYDEHPAKGLSRDEEEMALYSEWLVHGKSYNGLGEKDKREIFKDNLRIDEQNSRDRSYKLGLENFADDTNEYRSTYLGAHGTDARRRIKTKSDRRYAPKAGSLPDSIDMEKRGVAVKRGQSGSCAFSTIAVEGINOITVGLSLSBOELVDVDTSENGNGIMDYAFETIKNGGCTDEPDTPTTGRGCDTRKNAKNSIDYEDVTPFDEALAEVAGQSVATLRAAGRFOLYTSSTGIFGSGCTGDLHDGVTAVGTENGVDVTVIVKNSMAWSWEGRYLRAQNNVNDKNGLCIAIEPSYPTKTGENPMPGPSPPSPMPCDDYDECPTSTGCCVFPPEYCHFCAMGCSPLSESDHSCPHDYPCVHVSQGCPCPMKNSPGLVKRPMRRTPAKKIRNNGSSS"		
BASE COUNT	455 a 397 c 460 g 439 t		
ORIGIN			
Query Match	13.4%; Score 155.4; DB 8; Length 1751;		
Best Local Similarity	52.7%; Pred. No. 1.4e-29;		
Matches	476; Conservative 0; Mismatches 391; Indels 36; Gaps 5;		
DB	182 CGAAGAAGAGCGCAAGAGACTTGACATTTTTCACGATACTGAACTATATCAGGACAT	241	
DB	219 CGGAGAAGAGGATTAAGAGTTTGAATCTTCAAGATATCTCAGGTACATCGAGAGACA	278	
DB	242 GAATCAAAACAAATAATCACCCATCTTCATGCTTAGAGTGACCAAGTTGCTGACAT	301	
DB	279 G-----AACACCGCGGAGACCGGCTCTCAACAGCTCGAGCTGAACAGATTGCTGATCT	332	
DB	302 CACTCTCTCAAGATTTCAGCAAAAGTACTTGCAGACTCCCAAGAGATGTGTCGACCAAT	361	
DB	333 CACCAACGAGGAGTACCGGCTCAACCTACTCTGCGGCTTAAGACCGATGCTCGCGGAGAT	392	
DB	362 CAAATATGCGCAACAGAAATATGAAGAAGAACAAATATTTCTTGACACATCCACCTGCATC	421	
DB	393 CGCTAAGACGAAACGATCGGCGGTATGCTCCCAAGCGCGGCGACGTTGCTCGGACTC	452	
DB	422 ATGGATTGAGAGAAAAAGTGTATCATACCCCAATTAAGTCCAAAGGAGCGTGTGGAAG	481	
DB	453 CATTTAGCTGAGACGAAAGAGTGTGCGTCCGCAATCATCAAGAGATCTTGTGGAG	512	

OY 482 GGGTGTGGCGTTTCTGCGACGGGAGCCATGGAACACACATGCATTAATGAAACGAGGA 541
 || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 Db 513 TTGCTGGGCAATTCTAACCATTTGCTGCTGTGGAAAGCTATAAACCAAATCGTACTGGTGA 572
 OY 542 CCTTTTAGACCTTTCTGGAACAAGAACCTGTAGACAGTG---TGGAAAGAAGCAAGTTTC 598
 || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 Db 573 ATTGATCTCATTGTCTGAACAGAGACCTTGTTGGATTGTATACATCATTAACGAAGCGTGG 632
 OY 599 TTACATGAGGAGGCAATATCAATCGTTGCAATGGGTTTTAGAACATGGTGGCATGCCAC 658
 || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 Db 633 CAATGGTGGTCAATAGGAGCTATGCTTGGAGTTCAATCAAGAATGGTGGATTGATAC 692
 OY 659 TGATCATGATTATTCCTTCACAGAGCTAAAGAGGCTGATTCGAAGCCATTAGATACAGA 718
 || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 Db 693 TGAAGCAATTAACCCCTCACACTGAGTAGGTATGGAAGGTGATCAAAACGAAGAAAAATGC 752
 OY 719 CAAAG---GTTACAAATTGACGATATGAACACTTAATATGTCAGATGAGACAGAATC 775
 || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 Db 753 TAGAGTCGTTCTATTATGAGGTATGAGAGATGTTACTCCTATATGATGAGGGGCACTAAA 812
 OY 776 AGAGACAGAGCAAGCGTTCTTAAGCGGCATCCCTTGAGAACCAATTAAGTGTCTCAATTGA 835
 || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 Db 813 AGAGCGCAGT-----TGACAGGCCAACCAAGTATGATGTGGCATTGAAGCTGG 857
 OY 836 TGCAAAAGATTTTCATTATATACCGGGGGAGATTATGATGAGAAAACCTGACAGATCC 895
 || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 Db 913 ----GACCTGGAGCACAGGAGTGACAGCTGGTGGATATGATACAGAAAATGGTGTGATTA 968
 OY 956 CTGATATAGGAAAAATCATATGGGGAGAAAGATTGGGGAGAGATGCTTACATTTGGATCCA 1015
 || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 Db 969 TTGGATTGTTAAGAAGCTCATGCGCTGCATCCGGGAGAGAAAGCGCTATTTGAGAGATCCA 1028
 OY 1016 AAGAACAACGGGTATATTATTAGAGAGTGTGGTGGATGAATATTTCGCTCATATCCCAAC 1075
 || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 Db 1029 GCGTATGTGAAGAACAATAATGGTGTGTGTGATATCCGATAGAGCCCTTTTACCAGAC 1088
 OY 1076 CAA 1078
 || |
 Db 1089 CAA 1091

```

RESULT 12
LOCUS   AF454957                      1699 bp      mRNA       linear     PLN 14-JAN-2002
DEFINITION   Brassica oleracea senescence-associated cysteine protease (CP2)
ACCESSION   AF454957
VERSION     AF454957.1
KEYWORDS    .
SOURCE      Brassica oleracea.
ORGANISM    Brassica oleracea; Streptophyta; Embryophyta; Tracheophyta; Eukaryota; Viridiplantae; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE   1 (bases 1 to 1699)
AUTHORS    Coupe,S.A., Sinclair,B.K., Watson,L.M. and Eason,J.R.
TITLE       Cysteine proteases and broccoli senescence: cloning, characterization and contribution to the process
FEATURES             Source
Journal            2 (bases 1 to 1699)
Reference          Coupe,S.A., Sinclair,B.K., Watson,L.M. and Eason,J.R.
Authors            Direct Submission
Title              Submitted (05-DEC-2001) Crop and Food Research, Private Bag 11 600, Palmerston North 5301, New Zealand
Journal            Location/Qualifiers
Features            1..1699
Source              /organism="Brassica oleracea"
                    /db_xref="taxon:3712"
                    1..1699
gene

```

CDS

```

/gene="CP2"
25..1407
/gene="CP2"
/note="BocP2; similar to Arabidopsis thaliana cysteine
protease component of a protease-inhibitor complex;
expressed during senescence"
/codon_start=1
/product="senescence-associated cysteine protease"
/protein_id="AAL60579.1"
/db_xref="gi:18141283"
/translat="MGSVKVITLLAMIGVSAADMSIISYDEKHHITAEENERSDAE
VARIEAMEMEKHGKAOSNVGEERQRFEDNLRPFIDHNKNKLXYKILGLTRA
DITNEEFRIYLIGAKSKRKYLTSDRPORPRGDAIPDSVMRKGAIAAKYDOSCCS
CMASSTIGAVBIGIKRYTGDIISLSEDELVDCTSYOCCMGELIMADPFIETLNIGEI
DREEDTYPFKADGRCDOTDRNAKVTTIDAEEVENDEALKTLMADPISVAIEAGG
RAFOLYSGVFDDICGETHLGHVAAVGVEENGNDVIIVRSWGSGEBGYIKMANI
IAETPKCGIMEASYSIKKQNPNPGSPSPISKPTOCDKXSCPESENTCCCLFK
YGKCFPMWCCPLEPAATCCDDNTSCPHXEYVPCNGDICLSKNSPFSVKALKRPAPK
FMASRNINIA"
```

BASE COUNT	491 a	359 c	421 g	428 t
ORIGIN				
Query Match	13.4%	Score 155.2;	DB 8;	Length 1699;
Best Local Similarity	55.2%;	Pred. No. 1.5e-29;		
Matches 379;	Conservative 0;	Mismatches 283;	Indels 24;	Gaps 3;
Dy	401 TTGTACCATCACCCTGCATCATGGGATTGGAGAAAAAGGTGTCAACCAACTAA	460		
Dd	423 TGTGACGCGATATCTGTGATTCAGTTGACTGGAGAAAAGAGTGCCTGCCGCTTAA	482		
Oy	461 GTACCAAGGGGCTGTGGAAGGGTGTGGCGCTTCTCCACGGAGGACATAGAAGCAGC	520		
Dd	483 AGATCAGGAAGAGCTCGGGAGTGTGTGGGCATTTTCAACCATGGAGCGGTGGAGGAAT	542		
Oy	521 ACATCATATGACACAGAGAACCTTTGTAAGCTTTCTGAACAAGAACTGTAGACTGTGT	580		
Dd	543 AAACAAGATCGGACAGAGAACTTAATCTCTTGTCTAACAAGAACTGTGACTGTGA	602		
Oy	581 GGA---GAAAGCGAAGGCTTCTTCAAGAGATGGCAGATCAATCGTCGAATGGGTTTT	637		
Dd	603 CACATCTTRCAACACAGGGTGTGATATGAGAGTTTGTATGATTAATGCGTTGAATTCATAT	662		
Oy	638 AGAACATGGTGCAATGGCACCTAGTATGATTATTCCTTACAGAGCTAAAGAGGATAGTG	697		
Dd	663 TAAAAAGGTGGTATTGATTACGGAAGAGATTATTCATACAAAGTGTGTGACGCTTG	722		
Oy	698 CAAGCCATTAAGATACAGACAAAGTTTCAATTGACGGATATGAACCTTAATATATTC	757		
Dd	723 TGATCAGACCGAGAAAAACGCAAAAGTTGTCA-----CAATTGATGC	764		
Oy	758 AGATGAGAGTACAGATACAGAGACAGACAGCTTCTTAAAGCGCATCTGTGACCAAC	817		
Dd	765 CTATGAGGATGTGCTTAGAGAACACAGAGGCTGCGCTCAAGAAACTTTGGCTAACACCC	824		
Oy	818 AATTAGTGCTTCATTTGATGCAAAAAGATTTTCAATTATATAACCGGGGAATTTATGATGG	877		
Dd	825 TATTAGTGTCCGCAATCGAAGCTGGGGGCTCGGGCAATTCACAGTC--TATCTTCGGGTGT	881		
Oy	878 AGAAAACTGTCAATCTCGTATGGGATTAATTCACATTTTCTTACTTGGGTTATGTTTC	937		
Dd	882 ATTTATGGAATATGTGGAACAGACACTGAGCATGGAATTTGTAAGTGTGGGTTACGGAAAC	941		
Oy	938 AGCGGATGCTAGATTACATGATAGCGAAAAATTCATGGGAGAGAAAGATTGGGAGAAGA	997		
Dd	942 TGAGAAACGGAAGGACTACTGATGTTGAGAANAACATCATGGGAGAGAACTGGGAGAGAG	1001		
Oy	998 TGTGTACATTTTGGATCCAAAGAAACAGCGGTAATTTATTAGAGAGTGTGGATCAATTA	1057		
Dd	1002 CGGATACATAAAGATGGACGTAATTTGCGAGCCAACAGAAAGTGTGAATATGCCAT	1061		
Oy	1058 TTTTCCTTCATACCCACCAAAGAG	1083		
Dd	1062 GGAGGCTTCATACCCGATCAGAAAG	1087		

LOCUS	AR049578	1056 bp	DNA	linear	PAT 29-SEP-1999
DEFINITION	Sequence 1 from patent US 5824534.				
ACCESSION	AR049578				
VERSION	AR049578.1	GI:6005617			
KEYWORDS					
SOURCE	Unknown.				
ORGANISM	Unknown.				
REFERENCE	Unclassified.				
AUTHORS	1 (bases 1 to 1056)				
TITLE	Asano, M., Kawai, M., Miwa, T., and Nio, N. Aminopeptidase Cx, and a method of hydrolyzing a protein with the same				
JOURNAL	Patent: US 5824534-A 1 20-OCT-1998;				
FEATURES	Location/Qualifiers				
source	1..1056				
BASE COUNT	295 a 229 c 297 g 235 t				
ORIGIN					
Query Match	13.4%; Score 154.4; DB 6; Length 1056;				
Best Local Similarity	52.9%; Pred. NO. 2.6e-29;				
Matches	494; Conservative 0; Mismatches 401; Indels 39; Gaps 6;				
QY	103 CCAAGTTTACCCACACAGAAACAGTGTCTTCACTGTTCCAACTATGAAAGGTGACATG	162			
DB	38 CCACGTTGGCACACGAGGAGAGCTGATGTCCATGTACGAGCAGCGCTCGTGAAGCACG	97			
QY	163 GACGTTGTACCATTAACACGAGAAAGAGGCGCAAGAGACTTCAGATTTTCAAGATACT	222			
DB	98 GGAAGTGTACCAACGCGCTCGCGGAGAGAGAGAACCGCTTCAGATTTTCAAGACACAC	157			
QY	223 GCAACTATATCAAGGACATGATGCAAAACAGAAATCAACCCCATTTCTGATGTTAGAT	282			
DB	158 TCGCATATTCAGACACCACTCCGCGAGAGCCGAACCTACAAAGC-----TCGGAC	211			
QY	283 TGAACCAAGTTTCTGACATCACTCCTCAAGAGTTAGCAAAAAGTACTTGCAGTCCCA	342			
DB	212 TGAACCGTTCCTATCTCAACCAAGAGATTAAGGCGCCAAAGTCTTGGAAACCAAGA	271			
QY	343 AGGATGTGCGAGCAAAATCAAAATGGCCACAAAGAAATGAAAGGAACAAATATTCTT	402			
DB	272 TCGATCCCAACGCGAGCGTCCGAAAGACCCGAGCAACCGCTACGCCACAGC---TGTCG	328			
QY	403 GTGACCACTCACCTGCATCATGGATTTGGAGAAAAAAGGTGTCATCAACCAAGTAAAGT	462			
DB	329 GCGCAAAATTTGCTGATTCCTGTTGATTTGGAGAAAGAGGTGCTGCTCCTGTCAAAG	388			
QY	463 ACCAAGGGGGCTGTGAAGGGTGTGGCGTTTCTGCCACGGAGCCATTAGAAGCAGAC	522			
DB	389 ACCAAGGAGGCTGTGGAGCTGTGGGCAATTCACGAAATCGGTCAATAGAAGGAATAA	448			
QY	523 ATGCAATGCAACAGAGAGACTTGTAGCCTTTGCAACAAGAACTCTAGACTGTG---	579			
DB	449 ATTAGATATAGACAGGCAAGCATTTGTTATATGAAAGAAAGAAATGGTGGATTTGATA	508			
QY	580 TGGAAAGAAAGCAAGGTTCTTCAATGCAATGGATGGCAGATCAATCGTTGCAATGGGTTT	639			
DB	509 CTGGATATATACCAAGATATGCAATGGAGCACTTATGCAATATGCAATTTAGTTCAAT	568			
QY	640 AACATGGTGGCAATTCGCATGTATGATGATATCTTACAGAGCTAAAGAGGTGATGATCA	699			
DB	569 ACAATGGCGGCAATTTATTTGATGGAGATTAACCCATACCGTGGTGTGATGGTATGATCG	628			
QY	700 AAGCAATATAGATACAGACAAGC---GTTACATTTGACGAGATATGAAACTGTAATATGT	756			
DB	629 ACAATATATGAGAAATGCTAAGACGTCTTATGATGATGATACGAGATGATGCTCGCT	688			
QY	757 CAGATGAGAGTACGAATACAGACAGAGCAAGGCTCTTAAAGGCCATCTGTAGCAAC	816			

COMMENT
AJINOMOTO CO INC
OS None
OC Artificial sequences.
PN JP 1997294583-A/1
PD 16-NOV-1997
PE 14-FEB-1997 JP 1997030458
PR 08-MAR-1996 JP 96P 51848
PI ASANO MINAO, KAWAI MISAKO, MIWA TETSUYA, NIO NORIKI PC
C12N9/48/C12N15/00;
CC strandedness: Double;
CC topology: Linear;
FH key Location/Qualifiers
FT source 1..1056
FT FT /organism='Artificial gene'
FT CDS 1..1056
FT /product='Glycine max protease D3' FT
FT misc-feature 1..9 /note='vector-derived sequence'.
FEATURES
source Location/Qualifiers
1..1056
/organism='unidentified'
/db_xref='taxon:3264'
BASE COUNT 295 a 229 c 297 g 235 t
ORIGIN
Query Match 13.4%: Score 154.4: DB 6: Length 1056:
Best Local Similarity 52.9%: Pred. No. 2.6e-29;
Matches 494: Conservative 0: Mismatches 401: Indels 39: Gaps 6:
QY 103 CCAAGTTTACACACAGAAACAGAGTCTCTTCACACTGTTCCAACTATGGAAGTGAGCATG 162
DB 38 CCAAGTTTACACACAGAAACAGAGTCTCTTCACACTGTTCCAACTATGGAAGTGAGCATG 97
QY 163 GACGTCTCAACATTAACACAGAAAGAGGCAAGAGCTTGAGATTTCAGAAATTAAT 222
DB 98 GGAAGGTGTACAAAGCGCTCGGCGAGAGAGAGAGCGCTTCAGATCTCAAGACAAAC 157
QY 223 CGAATATATTCAGAGGACATGATGCAACAGAAATCACCCATTCATCGTTAGAT 282
DB 158 TCGGATTTCATCGACGACACCACTCCGCGAGAGACCGAACTTCAACG-----TCGGAC 211
QY 283 TGAACAAGTTTCTGCTGACATCTCTCAAGAGTTTCAGCAAAAAGTTACTTSCAAGCTCCCA 342
DB 212 TGAACCGGTTTGGTGTATCTACCAAGAGAAATACAGGCGCAAGTCTTGGAACCAAGA 271
QY 343 AGGATGTGCGCAAAATCAAAATGCGCAACAAAGAAATGAAGAAGAAATATTTCTT 402
DB 272 TCGATCCCAACCGAGGCTCGGAAGAACCCCGAGCAACCGCTACGCGCCACG--TGTG 328
QY 403 GTGACCATCCACTGCATCATGGGATTGGAGAAAAAGGTGTATCACCAGTAAGT 462
DB 329 GCGACAAATTTGCTGATTCGTTGAGGAGAAAGGTGCTGTCCTCCTGCAAG 388
QY 463 ACCAAGGGGCTGTGGAAGGGGTTGGGCGTTTCTGCCACGAGCCATTAAGACAGCAC 522
DB 389 ACCAAGAGGCTGTGGAGAGTGTGGGCAATTCACAGCAATCGGTGCAATGAAGAAATA 448
QY 523 ATGCAATAGCAAGAGACCTTGTAGCTTCTGAACAAGAACTGTAGACTGTG-- 579
DB 449 ATTAAGATAGTAACAGGCAATGATTTCTTATCAGAAACAAAGATTTGGATTGTGATA 508
QY 580 TGAAGAAGCGAAGGTCTTACAATGATGGAGATCAATGCTTCGAATGGGTTTAG 639
DB 509 CTGGATATACCAAGAGGACTTATGAGAGACTTATGCAATTTGAGTTGATCATATCA 568
QY 640 AACATGTTGGCATTTGCCACTGATGATATCTTACAGAGCTAAAGAGTAGATGCA 699
DB 569 ACAATGCGCGCATGATTCGATGAGAGATTACCATACCGTGTGTTGATGATGATGCG 628
QY 700 AAGCAATTAAGATACAAGCAAG--GTTACAATTGACGATATGAACCTTAATATGT 756
DB 629 ACACATATAGGAAAAATGCTAAAGTCTTTCTATGATGACTAGCAAGATGTTCTGCGCT 688

QY 757 CAGATGAGTATACAGATTCAGAGACAGAGCGTCTTAAGGCCATCTTGAGCAAC 816
DB 689 ATGATGAGTT-----AGCTTGAAGAAAGCCGTTGCAAAATCAGC 727
QY 817 CAATTAGTGTCTCAATTGATGACAAAGATTTTCATTATACACCGGGGAATTTATGATG 876
DB 728 CCGTGAAGCGTTGCTATTTGAAGAGGGGCGACAGCAATTCATATATATG---TATCTGCTG 784
QY 877 GAGAAACTGTACAACTCCGTATGCGATTATGCACTTTGTTTACTTGTGGTTATGTT 936
DB 785 TATTACGCGGAGAGATGCGCAGCACTAGATCATGCTGTGCTGTGGATAGAA 844
QY 937 CACCGGATGCTGATGATTTACTGATGACCAAAATTCATGCGGAGAGATTTGGGAGAG 996
DB 845 CAGCTAAAGGTCAATGATTTATTTGATCGTAAGAAATTCATGCGGTTCTAGCTGGGAGAGG 904
QY 997 ATGTTACATTTGATTCCAAGAAACACGCGGTAA 1030
DB 905 ATGCTCATCATGATTTAGAAAGAAATCTTCTTAA 938

Search completed: January 11, 2003, 20:37:37
Job time : 2682 secs

THIS PAGE BLANK (USPTO)


```

/tissue_type="Seed coats"
/lab_host="E. coli strain XL0LR"
/Note="Vector: pBK-CMV; Site.1: EcoRI; Site.2: XhoI; This
cDNA library was constructed from polyA+ enriched mRNA
from green seed coats in mid to late developmental stage
, average fresh weight 250 mg per seed. Traces of pod and
embryo tissue also present. Complementary DNA was
synthesized from mRNA using an XhoI-poly(dT)
linker-primer. EcoRI adapters were ligated to the
blunt-ended cDNA fragments and the products were digested
with XhoI for directional cloning into lambda Zap Express
vector. This lambda library was amplified once using E.
coli host strain XL1 Blue MRF+. Inserts were then
subcloned by mass excision using EXAssist helper phage for
conversion into phagemid vector pBK-CMV in E. coli host
strain XL0LR."
BASE COUNT      278 a      177 c      192 g      214 t      7 others
ORIGIN
Query Match      67.0%; Score 775; DB 10; Length 868;
Best Local Similarity 95.3%; Pred. No. 9.3e-207;
Matches 827; Conservative 0; Mismatches 37; Indels 4; Gaps 3;
QY 29 TCTTTTCTCCCTTTAGGCTCTCTTCTAGTCCAGCATATCAACTGATGTTCCATATT 88
DB 1 TCTTTTCTCCCTTTAGGCTCTCTTCTAGTCCAGCATATCAACTGATGTTCCATATT 60
QY 89 GGACCTTGACCTTAACCAAGTTTACACAGAAACAGGTCTCTTCACTGTTCCAACTATG 148
DB 61 GGACCTTGACCTTAACCAAGTTTACACAGAAACAGGTCTCTTCACTGTTCCAACTATG 120
QY 149 GAAGAGTGCAGCATGAGCGTGTCTACCAATACCAAGAAAGAGGCAAGACTTGAGAT 208
DB 121 GAAGAGTGCAGCATGAGCGTGTCTACCAATACCAAGAAAGAGGCAAGACTTGAGAT 180
QY 209 TTTCAGAAATTAATCTGCACTATATCAGGGACATGATGCAAAACAGAAATCACCATTTC 268
DB 181 TTTCAGAAATTAATCTGCACTATATCAGGGACATGATGCAAAACAGAAATCACCATTTC 240
QY 269 TCAATCGTTTGGAGTTGAACAAGTTTGGCTGACATCATCCTCCAGAGTTTCAGCAAAAAGTA 328
DB 241 TCAATCGTTTGGAGTTGAACAAGTTTGGCTGACATCATCCTCCAGAGTTTCAGCAAAAAGTA 300
QY 329 CTTCGCAAGCTCCCAAGAGATGTGTGCGACGAATCAAAATGGCCCAACAGAAATGAAGAA 388
DB 301 CTTCGCAAGCTCCCAAGAGATGTGTGCGACGAATCAAAATGGCCCAACAGAAATGAAGAA 360
QY 389 GGAACAATATTTCTTTGGACCATCACCCTGCATCATGGGATTGAGAAAAAGGTGTCTAT 448
DB 361 GGAACAATATTTCTTTGGACCATCACCCTGCATCATGGGATTGAGAAAAAGGTGTCTAT 420
QY 449 CACCAAGTAAAGTAAAGCAAGGGGCTGTGAAGGGGTTGGCGTTTCTGCCACGGGAGC 508
DB 421 CACCAAGTAAAGTAAAGCAAGGGGCTGTGAAGGGGTTGGCGTTTCTGCCACGGGAGC 480
QY 509 CATAGAAGCAGCAGCATGCAATATAGCAACAGAGACCTTGTAGGCTTCTGAACAAGAACT 568
DB 481 CATAGAAGCAGCAGCATGCAATATAGCAACAGAGACCTTGTAGGCTTCTGAACAAGAACT 540
QY 569 CGTAGACTGTGTGGAAGAAAGCAAGGTTTTCATATGATGGCAGTATCAATCTGTTGCA 628
DB 541 CGTAGACTGTGTGGAAGAAAGCAAGGTTTTCATATGATGGCAGTATCAATCTGTTGCA 600
QY 629 ATGGGTTTGAAGCAATGTTGGCAATGGCAGTATGATATTCCTTACAGAGATAAGA 688
DB 601 ATGGGTTTGAAGCAATGTTGGCAATGGCAGTATGATATTCCTTACAGAGATAAGA 660
QY 689 GGGTAGATGCAAAAGCCATTAAGATATCAAGCAAGGTTTACGAGGATGAGAACTCT 748
DB 661 NGGTAGATGCAAAAGCCATTAAGATATCAAGCAAGGTTTACGAGGATGAGAACTCT 720
QY 749 AATTAATGTCTGAG-TGAGAGTAGTACGAATCAGAGACAGAGGCTTCTTAAGGCCCA-TC 806

```

```

DB 721 TATAATGTCAATATTGAGAGTCCGGAATCAGAGCNGCAACAGCGTTCTTAACGCCATTTC 800
QY 807 CTTCGAGCAACCAATTTAGTGTCT--CAATTGATGCAAAAGATTTTCATTATACACGGGG 864
DB 781 CTTCGAGCAACCAATTTAGNGGCTTCCATTGATGCGCAAAAGATTTCTTTTCCCGGGG 840
QY 865 GAATTTATGATGAGAAAGTGTACAG 892
DB 841 GAATTTATGATGAGGGAAGAAACTTTTCCAG 868

RESULT 2
BE660308      845 bp      mRNA      linear      EST 06-sep-2000
LOCUS        BE660308      Glycine max cDNA, mRNA sequence.
DEFINITION   BE660308
ACCESSION    BE660308
VERSION      BE660308.1 GI:9986200
KEYWORDS     EST.
SOURCE       soybean.
ORGANISM     Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eustoids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.
REFERENCE    1 (bases 1 to 845)
AUTHORS      Harris,N., Chapman,B.P. and Gijzen,M.
TITLE        Gene expression in developing soybean seed coats
JOURNAL      Unpublished (2000)
COMMENT      Contact: Gijzen M
              Agriculture and Agri-Food Canada
              1391 Sandford Street, London, Ontario, Canada N5V 4T3
              Tel: 519 457 1470
              Fax: 519 457 3997
              Email: gijzenm@em.agr.ca

FEATURES
    source     location/Qualifiers
                1..845
                /organism="Glycine max"
                /cultivar="Harosoy 63"
                /db_xref="taxon:3847"
                /clone_lib="GmaxSC"
                /tissue_type="Seed coats"
                /lab_host="E. coli strain XL0LR"
                /note="Vector: pBK-CMV; Site.1: EcoRI; Site.2: XhoI; This
cDNA library was constructed from polyA+ enriched mRNA
from green seed coats in mid to late developmental stage
, average fresh weight 250 mg per seed. Traces of pod and
embryo tissue also present. Complementary DNA was
synthesized from mRNA using an XhoI-poly(dT)
linker-primer. EcoRI adapters were ligated to the
blunt-ended cDNA fragments and the products were digested
with XhoI for directional cloning into lambda Zap Express
vector. This lambda library was amplified once using E.
coli host strain XL1 Blue MRF+. Inserts were then
subcloned by mass excision using EXAssist helper phage for
conversion into phagemid vector pBK-CMV in E. coli host
strain XL0LR."
BASE COUNT      279 a      150 c      200 g      206 t      10 others
ORIGIN
Query Match      60.6%; Score 700; DB 10; Length 845;
Best Local Similarity 92.8%; Pred. No. 1.1e-185;
Matches 772; Conservative 0; Mismatches 55; Indels 5; Gaps 4;
QY 196 AGAGACTTGAGATTTTCAGAAATTAATCTGCAAGTATATCAGGGACATGAAATGCAAAAGAA 255
DB 1 AGAGACTTGAGATTTTCAGAAATTAATCTGCAAGTATATCAGGGACATGAAATGCAAAAGAA 60
QY 256 AATCACCCTTCCTCATCGTTAGATGTAACAAGTTTGGTGCATCATCCTCCAGAGAT 315
DB 61 AATCACCCTTCCTCATCGTTAGATGTAACAAGTTTGGTGCATCATCCTCCAGAGAT 120
QY 316 TCAGCAAAAAGTACTTGCAAGTCCCAAGAGATGTGTGCGACGAAATCAAAATGCGCCACA 375

```


Db 513 CGGTGGGCGCTGTTTCGCCACGGGAGCCATAGAAGCAGCATGCATAGCAACATGAGA 572

QY 542 CCTGTAGGCTTCTGAGACAGAACTCGTAGACTGTGTGAGAAAGCGAAGTCTCTTA 601
|||||

Db 573 CCTTGTAGGCTTCTGAGACAGAACTCGTAGACTGTGTGAGAAAGCGAAGTCTCTTA 632
|||||

QY 602 CAATGATGACAGTATCAATCGTTTGAATGGTTTGAACATGCTGGCACTGTA 661
|||||

Db 633 CAATGATGACAGTATCAATCGTTTGAATGGTTTGAACATGCTGGCACTGTA 692
|||||

QY 662 TGATGATATCCTTACAGACTAAAGAGGTAGTATGCAAGCCCAATAGATCAAGCAA 721
|||||

Db 693 TGATGATATCCTTACAGACTAAAGAGGTAGTATGCAAGCCCAATAGATCAAGCAA 752
|||||

QY 722 GGTT 725
|||||

Db 753 GGTT 756
|||||

RESULT 4
A1941099 729 bp mRNA linear EST 30-NOV-2001
LOCUS SB84409.y1 Gm-c1010 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:
DEFINITION Gm-c1010-1674 5' similar to TR:064458 064458 BD 30K.; mRNA
sequence.
ACCESSION A1941099
VERSION A1941099.1 GI:5688084
KEYWORDS EST.
SOURCE soybean.
ORGANISM Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.
REFERENCE 1 (bases 1 to 729)
AUTHORS Shoemaker,R., Keim,P., Vodkin,L., Erpelting,J., Coryell,V., Khanna
A., Bolla,B., Matra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,
Wylie,T., Underwood,R., Steptoe,M., Treising,B., Allen,M., Bowers
Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk
R., Riter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann
R., Waterston,R. and Wilson,R.
TITLE Public Soybean EST Project
JOURNAL Unpublished (1999)
COMMENT Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: estewatson.wustl.edu
Possible reversed clone: similarity on wrong strand This clone is
available through: ResGen, Invitrogen Corp. 2130 South Memorial
Parkway Huntsville, AL 35801 For further information call: (800
)-533-4363 or contact via email: ccu@resgen.com
High quality sequence stop: 382.
location/Qualifiers
1. 729
/organism="Glycine max"
/db_xref="taxon:3847"
/clone_lib="Gm-c1010"
/issue_type="young cotyledons of greenhouse grown plants"
/dev_stage="2cm long 12 week old"
/lab_host="XLI0-Gold"
/note="Vector: pBluescript II SK-; Site.1: EcoRI; Site.2:
XhoI; This cDNA library was constructed from mRNA isolated
from immature cotyledons (100-200mg) of old greenhouse
grown plants. The cDNA library was prepared using the
Stratagene pBluescript II SK(+) library construction kit.
Complementary DNA was synthesized from mRNA using a primer
consisting of a poly (dT) sequence with a XhoI restriction
site. EcoRI adapters were ligated to the blunt-ended cDNA
fragments followed by XhoI digestion. The cDNA fragments

BASE COUNT 228 a 158 c 169 g 173 t 1 others
ORIGIN
Query Match 59.7%; Score 690.2; DB 9; Length 729;
Best Local Similarity 97.4%; Pred. No. 6.3e-183;
Matches 701; Conservative 0; Mismatches 19; Indels 0; Gaps 0;
were directionally cloned into the EcoRI-XhoI restriction
site of the pBluescript vector. The ligated cDNA fragments
were transformed into XLI0-Gold host cells. This library
was constructed by Dr. Lilia Vodkin and Dr. Anu Khanna."

QY 18 CTTGTGTTGCTCTTCTTCTCCCTCTTAGGTCTCTCTTCAAGTTCACCATATCACTCAT 77
|||||

Db 10 CTTGTGTTGCTCTTCTTCTCCCTCTTAGGTCTCTCTTCAAGTTCACCATATCACTCAT 69
|||||

QY 78 CTTTCCATATGAGACCTTGACCTTAACCAATTTACACACAGAAACAGGTCTTCACTG 137
|||||

Db 70 CTTTCCATATGAGACCTTGACCTTAACCAATTTACACACAGAAACAGGTCTTCACTG 129
|||||

QY 138 TTCCAACTATGAGAGAGTGACATGACGTGTACCATTAACCAAGAGAGCAAG 197
|||||

Db 130 TTCCAACTATGAGAGAGTGACATGACGTGTACCATTAACCAAGAGAGCAAG 189
|||||

QY 198 AGACTTGAGATTTTCAAGATTAACCTGAACTATATCAGGACATGAATGCAAGAAAA 257
|||||

Db 190 AGACTTGAGATTTTCAAGATTAACCTGAACTATATCAGGACATGAATGCAAGAAAA 249
|||||

QY 258 TCACCCCATTCATCGTTTGAATGATTAACCAATTTGCTCACTCTCTCAAGAGTTC 317
|||||

Db 250 TCACCCCATTCATCGTTTGAATGATTAACCAATTTGCTCACTCTCTCAAGAGTTC 309
|||||

QY 318 AGCAAAAGTACTTTCAGAGCTCCCAAGAGATGTGTGAGCAAAATCAAAATGCGCAACAG 377
|||||

Db 310 AGCAAAAGTACTTTCAGAGCTCCCAAGAGATGTGTGAGCAAAATCAAAATGCGCAACAG 369
|||||

QY 378 AAAATGAAGAAGAACATATTTCTGTGACCATTCACCTCATATGAGATGAGAGAAA 437
|||||

Db 370 AAAATGAAGAAGAACATATTTCTGTGACCATTCACCTCATATGAGATGAGAGAAA 429
|||||

QY 438 AAAGGTCTCATCCCAAGAAAGTCCCAAGAGAGGAGGCTGTGAGAGGGGTTGGGTTTCT 497
|||||

Db 430 AAAGGTCTCATCCCAAGAAAGTCCCAAGAGAGGAGGCTGTGAGAGGGGTTGGGTTTCT 489
|||||

QY 498 GCCACGGAGCCATAGAGAGCAGCATGCAATAGCAACAGAGACCTTGTAGCCTTTCT 557
|||||

Db 490 GCCACGGAGCCATAGAGAGCAGCATGCAATAGCAACAGAGACCTTGTAGCCTATCT 549
|||||

QY 558 GAACAGAACTCGTAGACTGTGTGAGAAAGGAGGTTCTTACATGATGACAGTAT 617
|||||

Db 550 GAACAGAACTCGTAGACTGTGTGAGAAAGGAGGTTCTTACATGATGACAGTAT 609
|||||

QY 618 CAATCGTCAATGAGGTTTGAACATGATGATGATGATGATGATGATGATGATGATGAT 677
|||||

Db 610 CAATCGTCCGATGAGGTTTGAACATGATGATGATGATGATGATGATGATGATGATGAT 669
|||||

QY 678 AGAGCTTAAGAGGTAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 737
|||||

Db 670 AGAGCTTAAGAGGTAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 729
|||||

RESULT 5
A1941069 732 bp mRNA linear EST 30-NOV-2001
LOCUS SB84404.y1 Gm-c1010 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:
DEFINITION Gm-c1010-1639 5' similar to TR:064458 064458 BD 30K.; mRNA
sequence.
ACCESSION A1941069
VERSION A1941069.1 GI:5688054
KEYWORDS EST.
SOURCE soybean.
ORGANISM Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;


```

/clone="GENOME SYSTEMS CLONE ID: Gm-c1010-53"
/clone.lib="Gm-c1010"
/tissue_type="young cotyledons of greenhouse grown plants"
/dev_stage="2cm long 12 week old"
/lab_host="X110-Gold"
/note="Vector: pBluescript II SK+; Site_1: EcoRI; Site_2:
XhoI; This cDNA library was constructed from mRNA isolated
from immature cotyledons (100-200mg) of old greenhouse
grown plants. The cDNA library was prepared using the
Stratagene pBluescript II SK(+) library construction kit.
Complementary DNA was synthesized from mRNA using a primer
consisting of a poly (dT) sequence with a XhoI restriction
site. EcoRI adapters were ligated to the blunt-ended cDNA
fragments followed by XhoI digestion. The cDNA fragments
were directionally cloned into the EcoRI-XhoI restriction
site of the pBluescript vector. The ligated cDNA fragments
were transformed into X110-Gold host cells. This library
was constructed by Dr. Lila Vodkin and Dr. Anu Khanna."

```

```

BASE COUNT      220 a 150 c 156 g 177 t 2 others
ORIGIN
Query Match      57.1%; Score 659.8; DB 9; Length 705;
Best Local Similarity 97.9%; Pred No. 2.2e-174;
Matches 667; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
OY 9  AAGGGTTTCCTTGCTGCTCTTTCTCCCTTAGGTCCTCTTCTAGTCCAGCATA 68
    |||||||
DB 25  ATGGGTTTCCTTGCTGCTCTTTCTCCCTTAGGTCCTCTTCTAGTCCAGCATA 84
OY 69  TCACATCATCGTTCATATTGGACCTTGACCTAACCAAGTTTCCACACAGAAACAGGTG 128
    |||||||
DB 85  TCACATCATCGTTCATATTGGACCTTGACCTAACCAAGTTTCCACACAGAAACAGGTG 144
OY 129  TCTTCATGTTCCACTATGAGAGAGTGAAGCATGAGCTGTCTACCATACAGAGAA 188
    |||||||
DB 145  TCTTCATGTTCCACTATGAGAGAGTGAAGCATGAGCTGTCTACCATACAGAGAA 204
OY 189  GAGGCAAGAAGACTGAGATTTTCAAGAAATMACTGAACTATATCAAGGACATGAATGCA 248
    |||||||
DB 205  GAGGCAAGAAGACTGAGATTTTCAAGAAATMACTGAACTATATCAAGGACATGAATGCA 264
OY 249  AACAGAAATACACCCATCTCATCGTTTGAAGATTGAACAAGTTTGGTGCATCATCTCT 308
    |||||||
DB 265  AACAGAAATACACCCATCTCATCGTTTGAAGATTGAACAAGTTTGGTGCATCATCTCT 324
OY 309  CAAGAGTTCGCAAAAAGTACTTGAAGCTCCCAAGAGATGTGCGAGCAATCAAAATG 368
    |||||||
DB 325  CAAGAGTTCGCAAAAAGTACTTGAAGCTCCCAAGAGATGTGCGAGCAATCAAAATG 384
OY 369  GCCACAGAAAATGAAGAAGACAATATCTTGTGACCATCCACCTGCATCATGGAT 428
    |||||||
DB 385  GCCACAGAAAATGAAGAAGACAATATCTTGTGACCATCCACCTGCATCATGGAT 444
OY 429  TGGAGGAAAAAAGGTGTCTACACCAAGTATACCAAGGGGCTGTGGAAGGGTTGG 488
    |||||||
DB 445  TGGAGGAAAAAAGGTGTCTACACCAAGTATACCAAGGGGCTGTGGAAGGGTTGG 504
OY 489  GCGTTTTCGCCACGGGAGCATGAGAGCAATGACATATGACACAGAGACCTTGT 548
    |||||||
DB 505  GCGTTTTCGCCACGGGAGCATGAGAGCAATGACATATGACACAGAGACCTTGT 564
OY 549  AGCCTTTCGAAACAGAACTCTGAGACTGTGTGGAAGAAAGCAAGAGTCTTACATGGA 608
    |||||||
DB 565  AGCCTTTCGAAACAGAACTCTGAGACTGTGTGGAAGAAAGCAAGAGTCTTACATGGA 624
OY 609  TGGCACTATCATCTGTTCCGATGGTTTGAACAATGTGTGGCAATGCAATGAT 668
    |||||||
DB 625  TGGCACTATCATCTGTTCCGATGGTTTGAACAATGTGTGGCAATGCAATGAT 684
OY 669  TATCTTACAGAGCTAAAGAG 689
    |||||||
DB 685  TATCTTACAGAGCTAAAGAG 705

```

```

RESULT 7
AM472063
LOCUS
DEFINITION
AM472063 704 bp mRNA linear EST 03-DEC-2001
S119E07.Y1 Gm-c1029 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:
Gm-c1029-1693 5' similar to FR:064458 064458 BD 30K. ; mRNA
sequence.
ACCESSION
AM472063
VERSION
AM472063.1 GI:7042169
KEYWORDS
EST.
SOURCE
soybean.
ORGANISM
Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.
REFERENCE
1 (bases 1 to 704)
Shoemaker,R., Keim,P., Vodkin,L., Erpelting,J., Corryell,V., Khanna
,A., Bolla,B., Mairra,M., Hallier,L., Kucaba,T., Martin,J., Beck,C.,
Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers
,Y., Person,B., Swaller,T., Glibbons,M., Pape,D., Harvey,N., Schurk
,R., Rittler,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann
,R., Waterson,R. and Wilson,R.
Public Soybean EST Project
Unpublished (1999)
Contact: Shoemaker R/Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available through: Resgen, Invitrogen Corp. 2130
South Memorial Parkway Huntsville, AL 35801 For further information
call: (800)-533-4363 or contact via email: ccuteresgen.com
Insert Length: 1508 Std Error: 0.00
High quality sequence stop: 441.
FEATURES
source
1..704
/organism="Glycine max"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-c1029-1693"
/clone.lib="Gm-c1029"
/tissue_type="very young cotyledons of greenhouse grown
plants"
/lab_host="DH10B"
/note="Vector: pSPORT1; Site_1: Not I; Site_2: Sal I; This
cDNA library was constructed from mRNA isolated from very
young cotyledons (20-50mg fresh weight) of greenhouse
grown plants. The library was prepared using the Life
Technologies pSPORT1 cDNA library construction kit.
Complementary DNA was synthesized from mRNA using a
poly(dT) sequence with a Not I restriction site. Sal I
linker adapters were ligated to the blunt-ended cDNA
fragments followed by Not I digestion. The cDNA fragments
were directionally cloned into the Not I-Sal I restriction
site of the pSPORT1 vector. The ligated cDNA fragments
were transformed into E.coli Electromax DH10B host cells.
This library was constructed by Dr. Paul Keim and Dr.
Virginia Corryell."

```

```

BASE COUNT      220 a 148 c 155 g 180 t 1 others
ORIGIN
Query Match      56.7%; Score 655.8; DB 10; Length 704;
Best Local Similarity 97.4%; Pred. No. 2.9e-173;
Matches 666; Conservative 0; Mismatches 18; Indels 0; Gaps 0;
OY 9  AAGGGTTTCCTTGCTGCTCTTTCTCCCTTAGGTCCTCTTCTAGTCCAGCATA 68
    |||||||
DB 21  ATGGGTTTCCTTGCTGCTCTTTCTCCCTTAGGTCCTCTTCTAGTCCAGCATA 80
OY 69  TCACATCATCGTTCATATTGGACCTTGACCTAACCAAGTTTACCAACAGAAACAGGTG 128
    |||||||

```

Db 81 TCACATCATCGTTCCATATTGGACCTTGACCTAACCAGTTTACACAGAAACAGGTG 140
QY 129 TCCTCACTGTTCCAACTATGAGAGTGCATGAGAGCTGTCTACATTAACAGAGAA 188
Db 141 TCCTCACTGTTCCAACTATGAGAGTGCATGAGAGCTGTCTACATTAACAGAGAA 200
QY 189 GAGGCAAGAGACTTGATGATTTTCAGAAATTAATCGAATATATCAGGACATGAAATGCA 248
Db 201 GAGGCAAGAGACTTGATGATTTTCAGAAATTAATCGAATATATCAGGACATGAAATGCA 260
QY 249 AAGAGAAATACCCCATTTCTCATCGTTTAGATGGAACAAGTTGCTGACATCTCT 308
Db 261 AAGAGAAATACCCCATTTCTCATCGTTTAGATGGAACAAGTTGCTGACATCTCT 320
QY 309 CAAGAGTTCCAGAAAAGTCTTGCAAGCTCCCAAGATGTGTGCGACGAAATCAAAATG 368
Db 321 CAAGAGTTCCAGAAAAGTCTTGCAAGCTCCCAAGATGTGTGCGACGAAATCAAAATG 380
QY 369 GCCAACAGAAAATGAAGAAAGCAATATTCTTGACCATCCACCTGCATCATGGAT 428
Db 381 GCCAACAGAAAATGAAGAAAGCAATATTCTTGACCATCCACCTGCATCATGGAT 440
QY 429 TGGAGGAAAAAGGTGTGATCACCACCAAGTAAGTACCAAGGGGCTGTGAAAGGGTTGG 488
Db 441 TGGAGGAAAAAGGTGTGATCACCACCAAGTAAGTACCAAGGGGCTGTGAAAGGGTTGG 500
QY 489 GCGTTTCTGCCAGGGAGCCATAGAGACGACATCATATAGCAAGAGACCTTGT 548
Db 501 GCGTTTCTGCCAGGGAGCCATAGAGACGACATCATATAGCAAGAGACCTTGT 560
QY 549 AGCCTTCTGACAGAACTGTAGACTGTGTGGAAGAAAGCGAAGTTCTTACAAATGA 608
Db 561 AGCCTTCTGACAGAACTGTAGACTGTGTGGAAGAAAGCGAAGTTCTTACAAATGA 620
QY 609 TGGCAGATCATGCTGCGAATGGTTTAGAANCATGCTGGCATTCGCATGATGAT 668
Db 621 TGGCAGATCATGCTGCGAATGGTTTAGAANCATGCTGGCATTCGCATGATGAT 680
QY 669 TATCTTACAGAGCTAAAGAGGT 692
Db 681 TATCTTACAGAGCTAAAGATGT 704

RESULT 8
LOCUS BE660304 664 bp mRNA linear EST 06-SEP-2000
DEFINITION 2-ClD GmaxSC glycine max CDNA, mRNA sequence.
ACCESSION BE660304
VERSION BE660304.1 GI:9986196
KEYWORDS EST.
SOURCE soybean.
ORGANISM Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eustosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
glycine.
1 (bases 1 to 664)
Harris,N., Chapman,B.P. and Gijzen,M.
Gene expression in developing soybean seed coats
Unpublished (2000)
Contact: Gijzen M
Agriculture and Agri-Food Canada
1391 Sandford Street, London, Ontario, Canada N5V 4T3
Tel: 519 457 1470
Fax: 519 457 3997
Email: gijzenm@em.agr.ca.
Location/Qualifiers
1..664
/organism="Glycine max"
/cultivar="HaroSoy 63"
/db_xref="taxon:3847"
/clone_11b="GmaxSC"
/tissue_type="Seed coats"

/lab_host="E. coli strain XL0R"
/note="Vector: PBK-CMV; Site_1: EcoRI; Site_2: XhoI; This
cDNA library was constructed from polyA+ enriched mRNA
from green seed coats in mid to late developmental stage
, average fresh weight 250 mg per seed. Traces of pod and
embryo tissue also present. Complementary DNA was
synthesized from mRNA using an XhoI-poly(dT)
linker-primer. EcoRI adapters were ligated to the
blunt-ended cDNA fragments and the products were digested
with XhoI for directional cloning into lambda ZAP Express
vector. This lambda library was amplified once using E.
coli host strain Xl Blue MRF+. Inserts were then
subcloned by mass excision using EXSist helper phage for
conversion into phagemid vector PBK-CMV in E. coli host
strain XL0R."

BASE COUNT 223 a 132 c 155 g 151 t 3 others
ORIGIN

Query Match 56.2%; Score 649.8; DB 10; Length 664;
Best Local Similarity 98.5%; Pred. No. 1.4e-171;
Matches 654; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 67 TATCACTCATCGTTCCATATGAGACCTTGACCTTAACCAAGTTTACACAGAAACAGG 126
Db 1 TATCACTCATCGTTCCATATGAGACCTTGACCTTAACCAAGTTTACCAAGAAACAGG 60
QY 127 TGCTTCACTGTTCCACTATGAGAGTGAGCATGAGCGTGCTACCATTAACCAAGAG 186
Db 61 TGCTTCACTGTTCCACTATGAGAGTGAGCATGAGCGTGCTACCATTAACCAAGAG 120
QY 187 AAGAGCAAGAGACTGAGATTTTCAAGATAATACGAACTATATCAGGACATGATG 246
Db 121 AAGAGCAAGAGACTGAGATTTTCAAGATAATACGAACTATATCAGGACATGATG 180
QY 247 CAACACAAAATCAACCCATCTCATGTTAGATGGAACAATTTGCTGACATCTC 306
Db 181 CAACACAAAATCAACCCATCTCATGTTAGATGGAACAATTTGCTGACATCTC 240
QY 307 CTCAGAGTTCCAGAAAAGTCTGCAAGCTCCCAAGGATGTGTGCGACGAATCAAA 366
Db 241 CTCAGAGTTCCAGAAAAGTCTGCAAGCTCCCAAGGATGTGTGCGACGAATCAAA 300
QY 367 TGGCCAAAGAAAATGAGAGAGACAAATATCTGTGACCATCATCTGCATCATGGG 426
Db 301 TGGCCAAAGAAAATGAGAGAGACAAATATCTGTGACCATCATCTGCATCATGGG 360
QY 427 ATTGAGGAAAAAGGTGTATCAACCAAGTAAGTACCAAGGGGCTGTGGAAGGGTT 486
Db 361 ATTGAGGAAAAAGGTGTATCAACCAAGTAAGTACCAAGGGGCTGTGGAAGGGTT 420
QY 487 GGGGCTTTCGCCAGGGAGCCATAGAGACGACATGCAATAGCAAGAGACCTTG 546
Db 421 GGGGCTTTCGCCAGGGAGCCATAGAGACGACATGCAATAGCAAGAGACCTTG 480
QY 547 TTACCTTTTGAACAAGAACTCTAGACTGTGTGAAGAAGGAGGTTCTTCAATG 606
Db 481 TTACCTTTTGAACAAGAACTCTAGACTGTGTGAAGAAGGAGGTTCTTCAATG 540
QY 607 GATGCGATATCAATCGTTGCAATGGGTTTGAACAATGGTGCAATTCACATGATG 666
Db 541 GATGCGATATCAATCGTTGCAATGGGTTTGAACAATGGTGCAATTCACATGATG 600
QY 667 ATTATCTTACAGAGTAAGAGGGTGTGATGCAAGCAATTAAGATCAAGCAAGGTTA 726
Db 601 ATTATCTTACAGAGTAAGAGGGTGTGATGCAAGCAATTAAGATCAAGCAAGGTTA 660
QY 727 CAAT 730
Db 661 CAAT 664

RESULT 9
AW317682

LOCUS AM317682 659 bp mRNA linear EST 03-DEC-2001
 DEFINITION S956a06.y1 Gm-cl007 Glycine max cDNA clone GENOME SYSTEMS CLONE ID: Gm-cl007-395 5' similar to TR:064458 064458 BD 30K. ;, mRNA sequence.
 ACCESSION AM317682
 VERSION AM317682.1 GI:6747226
 KEYWORDS EST.
 SOURCE soybean.
 ORGANISM Glycine max
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
 REFERENCE 1 (bases 1 to 659)
 AUTHORS Shoemaker,R., Keim,P., Vodkin,L., Erpelting,J., Coryell,V., Khanna,A., Bolla,B., Marra,M., Hillier,L., Kueba,T., Martin,J., Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritzer,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and Wilson,R.
 TITLE Public Soybean EST Project
 JOURNAL Unpublished (1999)
 COMMENT Contact: Shoemaker R./Public Soybean EST Project
 Public Soybean EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 This clone is available through: ResGen, Invitrogen Corp. 2130 South Memorial Parkway Huntville, AL 35801 For further information call: (800)-533-4363 or contact via email: cou@resgen.com
 Seq primer: -40RP from Gibco
 High quality sequence stop: 451.
 FEATURES
 Location/Qualifiers
 1..659
 /organism="Glycine max"
 /db_xref="taxon:3847"
 /clone="GENOME SYSTEMS CLONE ID: Gm-cl007-395"
 /clone_1lb="Gm-cl007"
 /lab_host="DH10B"
 /note="Vector: pSPORT1; Site-1: SalI; Site-2: NotI; This cDNA library was constructed from mRNA isolated from immature cotyledons of greenhouse grown plants (individual seed fresh weight of 100-300mg). The library was prepared using the Life Technologies Superscript cDNA library construction kit. Complementary DNA was synthesized from mRNA using a poly (dT) sequence with a NotI restriction site. SalI linkers adapters were ligated to the blunt-ended cDNA fragments followed by NotI digestion. The cDNA fragments were directionally cloned into the NotI-SalI restriction site of the pSPORT1 vector. The ligated cDNA fragments were transformed into E. coli Electromax DH10B host cells. This library was constructed by Dr. Lila Vodkin and Dr. Anu Khanna."

BASE COUNT 218 a 124 c 157 g 159 t 1 others
 ORIGIN

Query Match 56.1%; Score 648.4; DB 10; Length 659;
 Best Local Similarity 98.9%; Pred. No. 3,5e-171;
 Matches 652; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

262 CCATTCATCGTGGTGGATTGACAAAGTTGCTGCATCACTCCCAAGAGTTGACGA 321
 |||||||
 Db 1 CCATTCATCGTGGTGGATTGACAAAGTTGCTGCATCACTCCCAAGAGTTGACGA 60
 322 AAAAGTACTGCAAGCTCCCAAGAGTGTGCGAGCAAAATCAAAATGCCCAAGAAAA 381
 |||||||
 Db 61 AAAAGTACTGCAAGCTCCCAAGAGTGTGCGAGCAAAATCAAAATGCCCAAGAAAA 120
 382 TGAAGAAGCAACATATCTTGTGACCATCCACTGCATATGAGATTGAGAGAAAAAG 441
 |||||||
 Db 121 TGAAGAAGCAACATATCTTGTGACCATCCACTGCATATGAGATTGAGAGAAAAAG 180

QY 442 GTGTCAACCCCAAGTAAGTACCAAGGGGGCTGTGGAAGGGTGGCGTTTTCGCCA 501
 |||||||
 Db 181 GTGTCAACCCCAAGTAAGTACCAAGGGGGCTGTGGAAGGGTGGCGTTTTCGCCA 240
 QY 502 CGGAGCCATAGAGAGCAGCATGCAATAGCAACAGAGACCTTGTAGCCTTTCGAA 561
 |||||||
 Db 241 CGGAGCCATAGAGAGCAGCATGCAATAGCAACAGAGACCTTGTAGCCTTTCGAA 300
 QY 562 AAGAAGCTGAGACTGTGTGGAAGAAAGCAAGCTTCTCAATGATGGCATATCAAT 621
 |||||||
 Db 301 AAGAAGCTGAGACTGTGTGGAAGAAAGCAAGCTTCTCAATGATGGCATATCAAT 360
 QY 622 CGTTGCAATGGTTTAGAACATGATGCGCATGTCCTGATGATGATATCCTTACAG 681
 |||||||
 Db 361 CGTTGCAATGGTTTAGAACATGATGCGCATGTCCTGATGATGATATCCTTACAG 420
 QY 682 CTAAGAGGTAGATGCAAGCCCAATAGATACAGACAAGCTTACAAATGACGATATG 741
 |||||||
 Db 421 CTAAGAGGTAGATGCAAGCCCAATAGATACAGACAAGCTTACAAATGACGATATG 480
 QY 742 AAATCTAATATGTCAGATGAGAGTACAGATCAGAGACAGCAAGCTTTAAGCG 801
 |||||||
 Db 481 AAATCTAATATGTCAGATGAGAGTACAGATCAGAGACAGCAAGCTTTAAGCG 540
 QY 802 CCATCCCTTGAGCAACCAATAGTGTCCATGATGATGCAAAAGTTTCAATTTACACG 861
 |||||||
 Db 541 CCATCCCTTGAGCAACCAATAGTGTCCATGATGATGCAAAAGTTTCAATTTACACG 600
 QY 862 GGGGAATTTATGATGGAAGAAAGCTGACAGTCCGATGGGATTAATCACTTTGTTTA 920
 |||||||
 Db 601 GGGGAATTTATGATGGAAGAAAGCTGACAGTCCGATGGGATTAATCACTTTGTTTA 659

RESULT 10

LOCUS AM317681 657 bp mRNA linear EST 03-DEC-2001
 DEFINITION S956a05.y1 Gm-cl007 Glycine max cDNA clone GENOME SYSTEMS CLONE ID: Gm-cl007-393 5' similar to TR:064458 064458 BD 30K. ;, mRNA sequence.

ACCESSION AM317681
 VERSION AM317681.1 GI:6747225
 KEYWORDS EST.
 SOURCE soybean.
 ORGANISM Glycine max

REFERENCE 1 (bases 1 to 657)
 AUTHORS Shoemaker,R., Keim,P., Vodkin,L., Erpelting,J., Coryell,V., Khanna,A., Bolla,B., Marra,M., Hillier,L., Kueba,T., Martin,J., Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritzer,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and Wilson,R.
 TITLE Public Soybean EST Project
 JOURNAL Unpublished (1999)
 COMMENT Contact: Shoemaker R./Public Soybean EST Project
 Public Soybean EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 This clone is available through: ResGen, Invitrogen Corp. 2130 South Memorial Parkway Huntville, AL 35801 For further information call: (800)-533-4363 or contact via email: cou@resgen.com
 Seq primer: -40RP from Gibco
 High quality sequence stop: 448.
 FEATURES
 Location/Qualifiers
 1..657
 /organism="Glycine max"


```

/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-cl007-393"
/clone_id="Gm-cl007"
/lab_host="DH10B"
/notes="Vector: pSPOR1; Site_1: SalI; Site_2: NotI; This
cDNA library was constructed from mRNA isolated from
Immature cotyledons of greenhouse grown plants (individuals
seed fresh weight of 100-300mg). The library was prepared
using the Life Technologies pSPscript cDNA library
construction kit. Complementary DNA was synthesized from
mRNA using a poly (dT) sequence with a NotI restriction
site. SalI linkers adapters were ligated to the
blunt-ended cDNA fragments followed by NotI digestion. The
cDNA fragments were directionally cloned into the
NotI-SalI restriction site of the pSPOR1 vector. The
ligated cDNA fragments were transformed into E. coli
Electromax DH10B host cells. This library was constructed
by Dr. Lila Voklin and Dr. Anu Khanna."

```

Query Match	Similarity	98.8%	Pred.	No. 1.6e-170	Length	657:
Best Local	Similarity	98.8%	Pred.	No. 1.6e-170	Length	657:
Matches	649:	Conservative	0:	Mismatches	8:	Indels
						Gaps
						0:
QY	129	TCCTCACTGTTTCCAACTATGGAAGATGAGTGAACGTGACGTCTACCATTAACCCAGAGAA	188			
Db	1	TCTTCACTGTTTCCAACTATGGAAGATGAGTGAACGTGACGTCTACCATTAACCCAGAGAA	60			
OY	189	GAGGCAAAAGAGACTTGACATTTTCAAGATTAATCTGAACATATATCAGGACATGAAATGCA	248			
Db	61	GAGGCAAAAGAGACTTGACATTTTCAAGATTAATCTGAACATATATCAGGACATGAAATGCA	120			
OY	249	AACACAAAATCCCCCATCTTCATCGTTAGATTGAACAAATTTTGCTACATCTCT	308			
Db	121	AACACAAAATCCCCCATCTTCATCGTTAGATTGAACAAATTTTGCTACATCTCT	180			
OY	309	CAGAAGTTCAGCAAAAAGTACTTGCAAGCTCCCAAGATGTGTCCAGCAAAATCAAAATG	368			
Db	181	CAGAAGTTCAGCAAAAAGTACTTGCAAGCTCCCAAGATGTGTCCAGCAAAATCAAAATG	240			
OY	369	GCCAAACAAGAAAATGAGAAAGAACAAATATCTTGACATCCACTGTCATCATGGAT	428			
Db	241	GCCAAACAAGAAAATGAGAAAGAACAAATATCTTGACATCCACTGTCATCATGGAT	300			
OY	429	TGGAGAAAAAAGTGTCTATCACCAAGTAAAGTACCAAGGGGGCTGTGGAAGGGGTTGG	488			
Db	301	TGGAGAAAAAAGTGTCTATCACCAAGTAAAGTACCAAGGGGGCTGTGGAAGGGGTTGG	360			
OY	489	GGTTTTCTGCCACGGGAGCCATTAACGACGACATGCAATAGCAACAGAGAACTTGT	548			
Db	361	GGTTTTCTGCCACGGGAGCCATTAACGACGACATGCAATAGCAACAGAGAACTTGT	420			
OY	549	AGCCTTTCTGAACAAGAACTCGTAGACTGTGTGAAGAAAGGAGAGTTCCTTAATGA	608			
Db	421	AGCCTTTCTGAACAAGAACTCGTAGACTGTGTGAAGAAAGGAGAGTTCCTTAATGA	480			
OY	609	TGGCAGTATCAATCGTTCGAATGGGTTTGAACAATGGTGCCATTGCCACTGATGAT	668			
Db	481	TGGCAGTATCAATCGTTCGAATGGGTTTGAACAATGGTGCCATTGATGAT	540			
OY	669	TATCTTACAGAGCTTAAGAGGGGTAGATGCCAAGCCAAATAGATACCAAGAGTTTACA	728			
Db	541	TATCTTACAGAGCTTAAGAGGGGTAGATGCCAAGCCAAATAGATACCAAGAGTTTACA	600			
OY	729	ATTGACGGATGAAGAACTCTATATATGTGATGAGATGAGACAGATTCAGAGACAG	785			
Db	601	ATTGACGGATGAAGAACTCTATATATGTGATGAGATGAGACAGATTCAGAGACAG	657			
RESULT	11					
LOCUS	AM433000	655 bp	mRNA	linear	EST	03-DEC-2001

DEFINITION	Accession	Version	Keywords	Source	Organism
sl103b1.1.y1 Gm-cl029 Glycine max cDNA clone	GM-cl029-118 5'				
Sequence.	AA433000				
EST.	AA433000.1	GI:6964307			
soybean.					
Glycine max					
Eukaryote: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;					
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;					
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;					
Glycine.					
1 (bases 1 to 655)					
Shoemaker,R., Kaim,P., Vodkin,L., Erpellding,J., Coryell,V., Khanna					
,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,					
Wylie,T., Underwood,K., Stepien,M., Theising,B., Allen,M., Bowers					
,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk					
,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann					
,R., Waterston,R. and Wilson,R.					
Public Soybean EST Project					
Unpublished (1999)					
Contact: Shoemaker R/Public Soybean EST Project					
Washington University School of Medicine					
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA					
Tel: 314 286 1800					
Fax: 314 286 1810					
Email: est@watson.wustl.edu					
This clone is available through: Resgen, Invitrogen Corp. 2130					
South Memorial Parkway Huntsville, AL 35801 For further information					
call: (800)-533-4363 or contact via email: ccluresgen.com					
Insert Length: 1190 Std Error: 0.00					
High quality sequence stop: 452.					
Location/Qualifiers					
1. 655					
/organism="Glycine max"					
/db_xref="taxon:3847"					
/clone="GENOME SYSTEMS CLONE ID: Gm-cl029-118"					
/clone_11b="Gm-cl029"					
/tissue_type="very young cotyledons of greenhouse grown plants"					
/lab_host="DH10B"					
/note="Vector: pSPORT1; Site_1: Not I; Site_2: Sal I; This					
cDNA library was constructed from mRNA isolated from very					
young cotyledons (20-50mg fresh weight) of greenhouse					
grown plants. The library was prepared using the Life					
Technologies pSuperscript cDNA library construction kit.					
Complementary DNA was synthesized from mRNA using a					
poly(dT) sequence with a Not I restriction site. Sal I					
linker adapters were ligated to the blunt-ended cDNA					
fragments followed by Not I digestion. The cDNA fragments					
were directionally cloned into the Not I-Sal I restriction					
site of the pSPORT1 vector. The ligated cDNA fragments					
were transformed into E.coli Electromax DH10B host cells.					
This library was constructed by Dr. Paul Kaim and Dr.					
Virginia Coryell."					
BASE COUNT	226 a	122 c	156 g	150 t	1 others
ORIGIN					
Query Match	55.0%	Score 635.4	DB 10	Length 655	
Best Local Similarity	98.2%	Pred. NO. 1.6e-167			
Matches 642; Conservative	0;	Mismatches 12;	Indels 0;	Gaps 0;	
194	AAAGAGCTTGAGATTTCAGAAATACGACATATACAGGGACATGAAATGCAAAACG	253			
1	AAAGGACTTGAGATTTCAGAAATACGACATATACAGGGACATGAAATGCAAAACG	60			
254	AAATACCCCATCTCATCTGCTTTAGATTGAAACAAAGTTTCTGCATCACTCTCTCAAGA	313			
61	AAAATCACCACCATCTCATCTGCTTTAGATTGAAACAAAGTTTCTGCATCACTCTCTCAAGA	120			
314	GTTGACGAAAAAGTATCTTGCAAGCTCCCAAGAGATGTGTGCGACGCAAAATCAAAATGGCCAA	373			

Db	121	GTTGACGAAAAAGTACTCTTGCAAGCTCCCAAGAGATGTGTGGCAGAACAATCAAAATGCGCCA	180
Oy	374	CAGAAAATGAAGAAGAACATAATTCTTTGTGACCATCCACCTGCATATGGGATTTGGAG	433
Db	181	CAGAAAATGAAGAAGAACATAATTCTTTGTGACCATCCACCTGCATATGGGATTTGGAG	240
Oy	434	GA AAAAAGTGTCATCACCCCAAGTAAATTTACCAAGGGGGCTGTGGAGGGGTGGCGCTT	493
Db	241	GA AAAAAGTGTCATCACCCCAAGTAAATTTACCAAGGGGGCTGTGGAGGGGTGGCGCTT	300
Oy	494	TTTCTCCACGGGAGCCATFAGAAGCAGCACATGCCAATFAGCAACAGAGACCTTTGTAGCCT	553
Db	301	TTTCTCCACGGGAGCCATFAGAAGCAGCACATGCCAATFAGCAACAGAGACCTTTGTAGCCT	360
Oy	554	TTTCTCAACAGAACCTCGTAGACTGTGTGGAGAAAGCGAAGGTTCTTCAATGATGGCA	613
Db	361	TTTCTCAACAGAACCTCGTAGACTGTGTGGAGAAAGCGAAGGTTCTTCAATGATGGCA	420
Oy	614	GTATCAATCGTTTGAATGGTTTTAGAACATGGTGGCATTTGGCACTGATGATGATATCC	673
Db	421	CTATCAATCGTTTGAATGGTTTTAGAACATGGTGGCATTTGGCACTGATGATGATATCC	480
Oy	674	TTACGAGACTTAAAGGGGTAGATGCAAAAGCCATAAGATACAAAGCAGGTTACATTTGA	733
Db	481	TTACGAGACTTAAAGGGGTAGATGCAAAAGCCATAAGATACAAAGCAGGTTACATTTGA	540
Oy	734	CGGATATGAAACTCTAAATTAATGTCTGATGAGAGTACAGATACAGACAGACAGCCGTT	793
Db	541	CTGATATGAAACTCTAAATTAATGTCTGATGAGAGTACAGATACAGACAGACAGCCGTT	600
Oy	794	CTTAAAGCCCATCTCTTGAGCAACCAATTAGTGTCTCAATTGATGCAAAAAGATTT	847
Db	601	CTTAAAGCCCATCTCTTGAGCAACCAATTAGTGTCTCAATTGATGCAAAAAGATTT	654
RESULT 12			
LOCUS	AM398014	657 bp	mRNA linear EST 03-DEC-2001
DEFINITION	sg11b107.y1 Gm-c1007 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:		
	Gm-c1007-1844 5' similar to TR:064458 064458 BD 30K.; mRNA sequence.		
ACCESSION	AM398014		
VERSION	AM398014.1	GI:6916484	
KEYWORDS	EST.		
SOURCE	soybean.		
ORGANISM	glycine max		
	Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
	Spermatophytes; Magnoliophyta; eudicotyledons; core eudicots;		
	Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;		
	Glycine.		
REFERENCE	1 (bases 1 to 657)		
AUTHORS	Shoemaker,R., Kaim,P., Vodka,L., Erpelidg,J., Coryell,V., Khanna		
	'A', Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,D., Beck,C.,		
	Wylie,T., Underwood,K., Stepike,M., Theising,B., Allen,M., Bowers		
	,Y., Person,B., Swaller,T., Gibbons.M., Pape,D., Harey,N., Schurk		
	,R., Ritter,E., Kohn,S., Ship,T., Jackson,Y., Cardenas,M., McCann		
	,R., Waterson.R. and Wilson.R.		
TITLE	Public Soybean EST Project		
JOURNAL	Unpublished (1999)		
COMMENT	Contact: Shoemaker R/Public Soybean EST Project Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA Tel.: 314 286 1800 Fax: 314 286 1810 Email: est@wustl.wustl.edu This clone is available through: ResGen, Invitrogen Corp., 2130 South Memorial Parkway Huntysville, AL 35801 For further information call: (800)-533-4363 or contact via email: cculeresgen.com Insert Length: 1074 Std Error: 0.00 Seq primer: -40RP from GlbCO High quality sequence stop: 437. Location/Qualifiers		
FEATURES			

```

source
    1..657
       /organism="Glycine max"
       /db_xref="taxon:3847"
       /clone="GENOME SYSTEMS CLONE ID: Gm-cl007-1844"
       /clone_id="Gm-cl007"
       /lab_host="DH10B"
       /note="Vector: pSPORT1; site.1: SalI; site.2: NotI; This cDNA library was constructed from mRNA isolated from immature cotyledons of greenhouse grown plants (individual seed fresh weight of 100-300mg). The library was prepared using the Life Technologies superscript cDNA library construction kit. Complementary DNA was synthesized from mRNA using a poly (dT) sequence with a NotI restriction site. SalI linkers adapters were ligated to the blunt-ended cDNA fragments followed by NotI digestion. The cDNA fragments were directionally cloned into the NotI-SalI restriction site of the pSPORT1 vector. The ligated cDNA fragments were transformed into E. coli Electromax DH10B host cells. This library was constructed by Dr. Lilia Vodka and Dr. Anu Khanna."

BASE COUNT      214 a   113 c   170 g   159 t     1 others
ORIGIN

Query Match          54.8%, Score 633.6; DB 10; Length 657;
Best Local Similarity 97.7%; Pred. No. 5e-167;
Matches 642; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

Oy  327 TACTTGCAGAGCTCCCAAGATGTGTGGACGAACAATGCCAACACAGAATAATGAG 386
      |||
Db   1 TACTTGCAAGCTCCCAGAGATGTGTGGACGAACAATCAAAATGGCCAACAGAAAATGAG 60

Oy  387 AAGAACAAATATTCTTTGTGACCATTCCACCTCATTGAGGATTGGAGAAAAAAGTGTTC 446
      |||
Db   61 AAGAACAAATATTCTTTGTGACCATTCCACCTCATGATGGAGTAAGAGAAAAAAGTGTTC 120

Oy  447 ATCACCGCAAGTAAATACCAAGGGGCTGTGGAAGGGTTGGGGCTTTTCGCCACGGGA 506
      |||
Db   121 ATCACCGCAAGTAAATACCAAGGGGCTGTGGAAGGGTTGGGGCTTTTCGCCACGGGA 180

Oy  507 GCCATAGAGACGACATCATGACATAGCAACAGAGACCTCTTGAAGCTTTCGAACAGAA 566
      |||
Db   181 GCCATAGAGACGACATCATGACATAGCAACAGAGACCTCTTGAAGCTTTCGAACAGAA 240

Oy  567 CTCGTAGACTGTGTGGAGAAAGCGAAGCTTTCATCATGATGGCAGATATCAATCGTTC 626
      |||
Db   241 CTCGTAGACTGTGTGGAGAAAGCGAAGCTTTCATCATGATGGCAGATATCAATCGTTC 300

Oy  627 GAATGGCTTTTAGACATGCTGGCCCTTCCCATGATGATTAATCCCTTACAGAGCTPAA 686
      |||
Db   301 GAATGGCTTTTAGACATGCTGGCCCTTCCCATGATGATTAATCCCTTACAGAGCTPAA 360

Oy  687 GAGGCTAGATGCAAAAGCCAAATAGATACAGAACAAGTGTTACAAATGACGATATGAAACT 746
      |||
Db   361 GAGGCTAGATGCAAAAGCCAAATAGATACAGAACAAGTGTTACAAATGACGATATGAAACT 420

Oy  747 CTAATTAATGTCAGATGACAGTACAGAAATCAGACAGCAAGSGTCTTAAAGGCGCATC 806
      |||
Db   421 CTAATTAATGTCAGATGACAGTACAGAAATCAGACAGCAAGSGTCTTAAAGGCGCATC 480

Oy  807 CTGGAACAACCAATTAAGTGTCTCAATTATGTCGAAAGATTTTCAATTATTAACACGGGGGA 866
      |||
Db   481 CTGGAACAACCAATTAAGTGTCTCAATTATGTCGAAAGATTTTCAANTTATTAACACGGGGGA 540

Oy  867 ATTATTAATGAGAAACCTGTACAAAGTCCGATATGGGATTAATCACTTTGTTTACTGTG 926
      |||
Db   541 ATTATTAATGAGAAACCTGTACAAAGTCCGATATGGGATTAATCACTTTGTTTACTGTG 600

Oy  927 GGTTATGGTACAGCGATGTGTAGATTACGTAGACGAAGCAAATAATCATTTGGGGAGAA 983
      |||
Db   601 GGTTATGGTACAGCGATGTGTAGATTACGTAGACGAAGCAAATAATCATTTGGGGAGAA 657

```

A1794978 665 bp mRNA linear EST 30-NOV-2001
 LOCUS sb74604.y1 Gm-c1010 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:
 DEFINITION Gm-c1010-727 5' similar to TR:064458 064458 BD 30K. ;, mRNA
 sequence.
 ACCESSION A1794978 GI:5342694
 VERSION A1794978.1
 KEYWORDS EST.
 SOURCE soybean.
 ORGANISM Glycine max
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
 Glycine.
 1 (bases 1 to 665)
 REFERENCE
 AUTHORS Shoemaker, R., Keim, P., Vodkin, L., Erpellding, J., Coryell, V., Khanna
 A., Bolla, B., Matra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C.,
 Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers
 Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk
 R., Ritter, E., Kohn, S., Shu, T., Jackson, Y., Cardenas, M., McCann
 R., Waterston, R. and Wilson, R.
 Public Soybean EST Project
 Unpublished (1999)
 COMMENT
 TITLE Contact: Shoemaker R/Public Soybean EST Project
 JOURNAL Public Soybean EST Project
 COMMENT Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 Possible reversed clone: similarity on wrong strand This clone is
 available through: Resgen, Invitrogen Corp. 2130 South Memorial
 Parkway Huntsville, AL 35801 For further information call: (800
)-533-4363 or contact via email: cc@resgen.com
 Insert Length: 1178 Std Error: 0.00
 High quality sequence stop: 400.
 Location/Qualifiers
 1. 665
 /organism="Glycine max"
 /db_xref="taxon:3847"
 /clone="GENOME SYSTEMS CLONE ID: Gm-c1010-727"
 /clone_lib="Gm-c1010"
 /tissue_type="young cotyledons of greenhouse grown plants"
 /dev_stage="2cm long 12 week old"
 /lab_host="XL10-Gold"
 /note="vector: pBluescript II SK-; Site 1: EcoRI; Site 2:
 XhoI; This cDNA library was constructed from mRNA isolated
 from immature cotyledons (100-200mg) of old greenhouse
 grown plants. The cDNA library was prepared using the
 Stratagene pBluescript II SK(+) library construction kit.
 Complementary DNA was synthesized from mRNA using a primer
 consisting of a poly (dT) sequence with a XhoI restriction
 site. EcoRI adapters were ligated to the blunt-ended cDNA
 fragments followed by XhoI digestion. The cDNA fragments
 were directionally cloned into the EcoRI-XhoI restriction
 site of the pBluescript vector. The ligated cDNA fragments
 were transformed into XL10-Gold host cells. This library
 was constructed by Dr. Lila Vodkin and Dr. Ann Khanna."
 BASE COUNT 214 a 130 c 157 g 161 t 3 others
 ORIGIN
 Query Match 54.5%; Score 630; DB 9; Length 665;
 Best Local Similarity 96.5%; Pred. No 5.2e-166;
 Matches 642; Conservative 0; Mismatches 23; Indels 0; Gaps 0;
 Oy 265 ATTCTCATGCTTTAGATGAGTGAACAAGTTTGTCTGACATCCTCTCAAGATTCAGCAAAA 324
 |||||||
 Db 1 ATTCTCATGCTTTAGATGAGTGAACAAGTTTGTCTGACATCCTCTCAAGATTCAGCAAAA 60
 |||||||
 Oy 325 AGTACTTGCAAGCTCCCAAGAGATGTGCGACGACAAATCAAAATGGCCACACAGAAATGA 384
 |||||||
 Db 61 AGTACTTGCAAGCTCCCAAGAGATGTGCGACGACAAATCAAAATGGCCACACAGAAATGA 120
 |||||||

Oy 385 AGAAGCAACATATTTCTTGTCGACCATCCACCTGCATCATGGATGGAGAAAAAGGTG 444
 |||||||
 Db 121 AGAAGCAACATATTTCTTGTCGACCATCCACCTGCATCATGGATGGAGAAAAAGGTG 180
 |||||||
 Oy 445 TCATCACCAGTAAGTATACCAAGGGGCTGTGCAAGGGGTGGCGCTTTCTCCACGG 504
 |||||||
 Db 181 TCATCACCAGTAAGTATACCAAGGGGCTGTGCAAGGGGTGGCGCTTTCTCCACGG 240
 |||||||
 Oy 505 GAGCCATAGAACGACGACATGCATATAGCAACAGAGACCTGTAGCCTTCTCAACAG 564
 |||||||
 Db 241 GAGCCATAGAACGACGACATGCATATAGCAACAGAGACCTGTAGCCTTCTCAACAG 300
 |||||||
 Oy 565 AACTCTAGACTGTGTGGAAAGAACGAGGTTCTTCAATGATGGCAGTATCAATCGT 624
 |||||||
 Db 301 AACTCTAGACTGTGTGGAAAGAACGAGGTTCTTCAATGATGGCAGTATCAATCGT 360
 |||||||
 Oy 625 TCGAATGGCTTTTGAACATGTGGCATGGCTGCTGATGATGATTCCTTACAGAGCTA 684
 |||||||
 Db 361 TCGAATGGCTTTTGAACATGTGGCATGGCTGCTGATGATGATTCCTTACAGAGCTA 420
 |||||||
 Oy 685 AAGAGGGTAGATGCAAAAGCCATTAAGATACAGACAGCAAGGTTACATGACGATATGAAA 744
 |||||||
 Db 421 AAGAGGGTAGATGCAAAAGCCATTAAGATACAGACAGCAAGGTTACATGACGATATGATA 480
 |||||||
 Oy 745 CTCTAATATGTGATGAGATGAGATGACAGATCAGACAGACAGCGTCTTAAAGCGCA 804
 |||||||
 Db 481 CTCTAATATGTGATGAGATGAGATGACAGATCAGACAGACAGCGTCTTAAAGCGCA 540
 |||||||
 Oy 805 TCCTTGAGCAACCAATAGTGTCTCATATGTATGCAAAAAGATTTTCAATTATACCGGGG 864
 |||||||
 Db 541 TNCCTTGAGCAACCAATAGTGTCTCTCATGTATGCAAAAAGATTTTCAATTATACCGCGNG 600
 |||||||
 Oy 865 GAATTTATGATGAGCAAAACATGCTACAGTCCGCTATGGATTAATCACTTTTACTTG 924
 |||||||
 Db 601 GCATCTATGATGAGCAAAACATGCTCCAGTCCGTATCGATTAATCACTGTCTACTTG 660
 |||||||
 Oy 925 TGGGT 929
 |||||
 Db 661 TNGGT 665
 |||||
 RESULT 14 654 bp mRNA linear EST 03-DEC-2001
 LOCUS AM397984
 DEFINITION sg70906.y1 Gm-c1007 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:
 Gm-c1007-1811 5' similar to TR:064458 064458 BD 30K. ;, mRNA
 sequence.
 ACCESSION AM397984 GI:6916454
 VERSION AM397984.1
 KEYWORDS EST.
 SOURCE soybean.
 ORGANISM Glycine max
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
 Glycine.
 1 (bases 1 to 654)
 REFERENCE
 AUTHORS Shoemaker, R., Keim, P., Vodkin, L., Erpellding, J., Coryell, V., Khanna
 A., Bolla, B., Matra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C.,
 Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers
 Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk
 R., Ritter, E., Kohn, S., Shu, T., Jackson, Y., Cardenas, M., McCann
 R., Waterston, R. and Wilson, R.
 Public Soybean EST Project
 Unpublished (1999)
 COMMENT
 TITLE Contact: Shoemaker R/Public Soybean EST Project
 JOURNAL Public Soybean EST Project
 COMMENT Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 This clone is available through: Resgen, Invitrogen Corp. 2130


```

OY 69 TCAACTATCGTCCATATTGGACCTTGACCTTAACCAAGTTTACCACACAGAAACAGGTG 128
    |||||||
Db 78 TCAACTCATCGTTCCATATTGGACCTTGACCTTAACCAAGTTTACCACACAGAAACAGGTG 137
    |||||||
OY 129 TCTTCACTGTTCCAAATATGGAAGATGAGATGAGCTGTCTACCATTAACCAAGAA 188
    |||||||
Db 138 TCTTCACTGTTCCAAATATGGAAGATGAGATGAGCTGTCTACCATTAACCAAGAA 197
    |||||||
OY 189 GAGGCAAGAGACTTGGATTTTCAAGATTAATCGAATATATCAGGACATGAATGCA 248
    |||||||
Db 198 GAGGCAAGAGACTTGGATTTTCAAGATTAATCGAATATATCAGGACATGAATGCA 257
    |||||||
OY 249 AACGAAAAATCACCCCATTTCTCATCGTTTAGATGAACAAGTTTGTGACATCATCTCT 308
    |||||||
Db 258 AACGAAAAATCACCCCATTTCTCATCGTTTAGATGAACAAGTTTGTGACATCATCTCT 317
    |||||||
OY 309 CAAGAGTTCAGCAAAAAGTACTTGCAAGCTCCCAAGATGTGTGCGACCAATCAAAATG 368
    |||||||
Db 318 CAAGAGTTCAGCAAAAAGTACTTGCAAGCTCCCAAGATGTGTGCGACCAATCAAAATG 377
    |||||||
OY 369 GCCAACAAGAAATGAGAAGAGAACAAATATTTCTGTGACATCCACCTGCATCATGGGAT 428
    |||||||
Db 378 GCCAACAAGAAATGAGAAGAGAACAAATATTTCTGTGACATCCACCTGCATCATGGGAT 437
    |||||||
OY 429 TGAAGAAAAAAGGTGTATCATCACCAAGTAAAGTACCAGGGGCTGTGAAGGGTTGG 488
    |||||||
Db 438 TGAAGAAAAAAGGTGTATCATCACCAAGTAAAGTACCAGGGGCTGTGAAGGGGATG 497
    |||||||
OY 489 GCGTTTCTGCGACGGAGCATGGAAGCAGCATATGCAATAGCAACAGAGACCTTGT 548
    |||||||
Db 498 GCGTTTCTGCGACGGAGCATGGAAGCAGCATATGCAATAGCAACAGAGACCTTGT 557
    |||||||
OY 549 AGCCTTCTGAACAAGAACTGTAGACTGTGTGGAAGAAAGCGAAGTTTCTTACAATGA 608
    |||||||
Db 558 AGCCTTCTGAACAAGAACTGTAGACTGTGTGGAAGAAAGCGAAGTTTCTTACAATGA 617
    |||||||
OY 609 TGGCAGTATCAATCGTTGGAATGGGTTTGTAGAACATGATGCGCATTTGCCACTGATGATGAT 668
    |||||||
Db 618 TGGCAGTATCAATCGTTGGAATGGGTTTGTAGAACATGATGCGCATTTGCCACTGATGATGAT 677
    |||||||

```

Search completed: January 11, 2003, 21:09:52
 -Job time : 1933 secs

THIS PAGE BLANK (USPTO)


```

:
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 1..1056
: OTHER INFORMATION: /note="METHOD OF DETERMINING THE
: OTHER INFORMATION: CHARACTERISTICS: P"
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: 1..9
: OTHER INFORMATION: /note="INSERTION SEQUENCE, METHOD
: OTHER INFORMATION: OF DETERMINING THE CHARACTERISTICS: E"
US-09-500-651-1

Query Match      13.4%; Score 154.4; DB 1; Length 1056;
Best Local Similarity 52.9%; Pred. No. 1e-38;
Matches 494; Conservative 0; Mismatches 401; Indels 39; Gaps 6;

QY 103 CCAAGTTTACACACAGAAACAGTGTCTTCCACTATGACAGTGCATG 162
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 38 CCAGTTGCCACCGAGGAGAGCTGATGCCATGACGAGCGTCTGTGAAGCAG 97
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 163 GAGGTGTCTACCAACACAGAGGCAAGAGACTGTGATTTTCAAGAACT 222
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 98 GGAAGGTGTACACCGCGCTGGCGAGAGAGAGAGCGCTTCCAGATCTCAAGACAC 137
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 223 CGAATATATCAGGACATGGAATGCAAAACGAAATACCCCATTTTCATCGTTAGAT 282
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 158 TGGGATTCATCGACGACACCACTCCGCGGAGAGACCGCAACTCAAGC-----TGGAC 211
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 283 TGAACAAGTTTGTGATCATCTCTCAAGATTCAGCAAAAAGTACTTGCAGAGCTCCA 342
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 212 TGAACGGTGTGCTGATCTCCACACAGAGGAATACAGGGCCAGTACTTGGGAACCAAGA 271
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 343 AGATGTGTGCGACGAATCAAAATGCGCAACAAAGAAATGAAGAGGAATATTTCT 402
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 272 TCATGCCAACCGGAGAGCTGGAAAGACCCCGACACCGCTTCAGGCCACG---TGTG 328
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 403 GTGACCATCCACCTGCATCATGGGATTGAGAGAAAAAGGTGTATCATCCCAAGTAAGT 462
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 329 GCGACAAATTTGCTGATTCGCTTGAATGGAGAGAGAGTGTCTCCCGTGTCAAG 388
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 463 ACCAAGGGGCTGTGAGAGGGGCTTTTCTGCGACGGAGCCATGAAAGCAC 522
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 389 ACCAAGGAGCTGTGGAGGCTGTGGGCATTCAGCAATCGGTGAGTGAAGAGATAA 448
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 523 ATCAATAGCAGAGAGACTGTGAGCTTTCTGTAACAGAGAGACTGTGAGCTGTG--- 579
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 449 ATAGATAGTAAACAGCGAATCTATTTCTTATACAGACAGAAATGTGGATTTGATA 508
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 580 TGAAGAAACGGAAGTCTTACATGATGAGCATATCAATCGTTTGAATGGGTTTAC 639
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 509 CTGATATATACCAAGGATGCAATGAGGACTTATGACATTCGATTTGATTCATATCA 568
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 640 AACATGTGTCATTCGCACATGATGATATCTTACAGAGCTTAAAGAGGTAGATGCA 699
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 569 ACATGCGCGCATTTGATTCGATGAGAGATTACCATACCTGTGTGTGATGATGATGCG 628
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 700 AACCAATAGATACAGACAG---GTACAAATGACGATATGAAACCTATATATGT 756
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 629 ACACATATAGGAAAAATGCTAAAGTGTCTTATGATGACTGCAAGATGTCTGCT 688
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 757 CAGATGAGATACAGATCAGACAGAGCAGACCGCTTTAAGCGCCATCTTGAGACAC 816
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 689 ATGATGAGTT-----ACCTTTGAAAAAGGCGGTTGCAAAATAGC 727
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 817 CAATTGTGTCTCAATTGATGCAAAAAGATTTTATTTATACACCGGGGCAATTTANGAT 876
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 728 CCGTGAAGCGTGTATTTGAAGAGGGGGGAGGAAATTTCAATTAATG---TATCTG 784
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 877 GAGAAAAGTGTACAGCGGTATGGGATTAATGACTTTGTTACTGTGGTATAGCTT 936
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 785 TATTTACGGGAGATGTGGCAGACACTAGATCATGTGTGCGCTGTGGTATAGGAA 844
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 937 CAGCGGATGTGTAGATTACTGTGATAGCGAAAAATTCATGGGAGAAAGATTGGGAGAG 966
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```

DB 845 CACCTAAGGTATGATTAATTTGATGCTGAAGAAATTCATGAGGCTTACTGGGAGAG 904
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 997 ATGTTACATTTTGCATCCAAAGAAACACCGCTAA 1030
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 905 ATGGCTACATCAGATTAGAAAGAAATCTTGCTAA 938
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 2
US-08-813-591-1
: Sequence 1, Application US/08813591
: Patent No. 5824534
: GENERAL INFORMATION:
: APPLICANT: ASANO, MINAO
: APPLICANT: KAWAI, MISAKO
: APPLICANT: MIRA, TETSUYA
: APPLICANT: NIO, NORIKI
: TITLE OF INVENTION: AMINOPEPTIDASE GX, AND A METHOD OF
: TITLE OF INVENTION: HYDROLYSING A PROTEIN WITH THE SAME
: NUMBER OF SEQUENCES: 8
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
: ADDRESSEE: P.C.
: STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
: CITY: ARLINGTON
: STATE: VA
: COUNTRY: USA
: ZIP: 22202
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: patentin Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/813,591
: FILING DATE: 07-MAR-1997
: CLASSIFICATION: 530
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: JP 051848/1996
: FILING DATE: 08-MAR-1996
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: JP 030458/1997
: FILING DATE: 14-FEB-1997
: ATTORNEY/AGENT INFORMATION:
: NAME: OBLON, NORMAN F.
: REGISTRATION NUMBER: 24,618
: REFERENCE/DOCKET NUMBER: 10-845-0
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 703-413-3000
: TELEFAX: 703-413-2220
: INFORMATION FOR SEQ. ID NO. 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1056 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA to mRNA
: ORIGINAL SOURCE:
: ORGANISM: Glycine max
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 1..1056
: OTHER INFORMATION: /note="METHOD OF DETERMINING THE
: OTHER INFORMATION: CHARACTERISTICS: P"
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: 1..9
: OTHER INFORMATION: /note="INSERTION SEQUENCE, METHOD
: OTHER INFORMATION: OF DETERMINING THE CHARACTERISTICS: E"
US-08-813-591-1

Query Match      13.4%; Score 154.4; DB 1; Length 1056;
Best Local Similarity 52.9%; Pred. No. 1e-38;
```


Matches 494; Conservative 0; Mismatches 401; Indels 39; Gaps 6;

```
OY 103 CCAAGTTTCCACACAGAACAGAGTGTCTTCACTGTTCACATGATGAGAGTACGACANG 162
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 38 CCAAGTTTGGCGCACCGGAGGAGAGCTATATCCATGTACAGAGCTGCTCGTAACGACG 97
OY 163 GAGGTGTCTACCATACACAGAAAGAGGCAAGAGACTTGATTTTTCAGAAATTAAGT 222
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 98 GGAAGGTATACAAACCGGCTCGGGAGAGAGAGAGCGCTCCGATCTTCAAGACCAACC 157
OY 223 CGAATATATTCAGGAGCATGATGCAAAACAGAAATACCCCATTTCTCATCGTTTAGAT 282
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 158 TCGATTTATTCAGACGACCACTCCGCGAGAGCCGAACTTACAAAGC-----TCGGAC 211
OY 283 TGAACAAGTTGTGTCATCATCTCCATGAGTTTCAAGAAAAAGTCTTGCAAGTCCCA 342
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 212 TGAACCGGTTCCGTGATCTTCACACAGAGAAATACAGGCGCAAGTACTTGGAACCAAG 271
OY 343 AGGATGTGTGCGAGCAAAATCAAAATGCGCACAAAGAAATGAAGAAGCAATATTTCT 402
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 272 TCGATCCCAACCGGAGAGCTCGGAAAGACCCGAGCAACCGCTACGCGCACG---TGTG 328
OY 403 GTGACCATTCACCTGCTATGAGGATGAGAGAAAAAGGTGTCTATCACCCTAAGTAAGT 462
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 329 GCGACAAATTTGCTGATTCCTGTGATGAGAGAGAGAGTGTCTTCCCTGTCTCAAG 388
OY 463 ACCAAGGGGCTGTGGAAGGGGTTGGGCGTTTCTCCACGCGGACCATATGAAAGCAGC 522
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 389 ACCAAGGAGGCTGTGGAGACTGTGGGCAATTCACAGCAATCGGTGACAGTAAGAAATTA 448
OY 523 ATGCATATACACAGAGAGACTTGTAGCCCTTCTGACAGAAAGACGTAAGACTG--- 579
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 449 ATAAAGATATTAACAGCGCAACTGATTTGCTTATCAGAACAGAAATGTTGGATTTGATA 508
OY 580 TGGAGAAAGCCGAGGCTTTTACATGATGCGAGTATGAATCGTTTCAAGGCGTTTATG 639
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 509 CTGGATATTAACCAAGATCAATGAGAGACTTATGACATATGATTTGATTCATATCA 568
OY 640 AACATGTCGCAATTCGCGATGATGATTTATTCCTTACAGAGCTAAAGGGATGATGCA 699
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 569 ACAATGGCGGCTATGATTTGATGAGATTAACCATACGCTGTTGATGATGATGATGCG 628
OY 700 AAGCCATATAGATACAAAGC---GTTACAAATGACGATATGAAGCTCTAATATGT 756
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 629 ACACATATAGGAAATATGCTAAAGCTTTCTATGATGATGATGATGATGATGATGATG 688
OY 757 CAGATAGAGTACAGATACAGACAGACAGACAGCGTTCTTAAGCGCCATCTTGAGCAAC 816
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 689 ATGATGAGTT-----AGCCTTGAAAAAAGCGCTTGCAAAATCAGC 727
OY 817 CAATGTGTCTCAATTCATGCAAAAGATTTTCATTTATACACCGGGGCAATTTATGATG 876
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 728 CCGTGAAGCTGTGATTTGAAGGAGGCGGCAAGGAAATTCATTAATATATG---TATCTG 784
OY 877 GAGAAAACTGTACAGTCCGATATGAGTTAATCACTTGTGTTTACTTGTGGTATGATG 936
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 785 TATTCACGGGAGATGTGACACAGACATGATCATGTGTCTGTGGTATAGGAA 844
OY 937 CAGCGGATGTGTAGATTCTGTGATAGCGAAAAATTCATGGGGAGAGATTTGGGGAGAA 996
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 845 CAGCTTAAAGGTATGATTTGATGATGATGATGATGATGATGATGATGATGATGATG 904
OY 997 ATGCTTACATTTGATTCACAAAGAAACAGCGGTTAA 1030
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 905 ATGCTTACATCAGATTAGAAAGAAATCTTGCTTA 938
```

RESULT 3
US-08-821-994-59
; Sequence 59, Application US/08821994A
; Patent No. 6228643
; GENERAL INFORMATION:
; APPLICANT: Greenland, Andrew J

```
; APPLICANT: Thomas, Didier RP  
; APPLICANT: Jepson, Ian  
; TITLE OF INVENTION: Promoters  
; FILE REFERENCE: PPD 50108  
; CURRENT APPLICATION NUMBER: US/08/821,994A  
; CURRENT FILING DATE: 1997-03-22  
; EARLIER APPLICATION NUMBER: PCT/GB97/00729  
; EARLIER FILING DATE: 1997-03-18  
; EARLIER APPLICATION NUMBER: GB 9606062.9  
; EARLIER FILING DATE: 1996-03-22  
; NUMBER OF SEQ ID NOS: 89  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 59  
; LENGTH: 1577  
; TYPE: DNA  
; ORGANISM: Brassica napus  
US-08-821-994-59
```

Query Match 13.1%; Score 151.2; DB 4; Length 1577;
Best Local Similarity 55.2%; Pred. No. 1.3e-37;
Matches 369; Conservative 1; Mismatches 274; Indels 24; Gaps 3;

```
OY 414 CCTGCATCATGGGATTTGAGAGAAAAAGGTGTCATACCCCAAGTAAAGTACCAAGGGGC 473
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 402 CCAGATTTCTGTGATTTGAGAGAAAGAGCTGTACTAATATGTCAAAGATCAAGAGAAC 461
OY 474 TGTGGAAGGGGTGGCGTTTCTGCCAGGAGACCATAGAGACACATGATATAGCA 533
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 462 TGGGAGCGGTGGTGTCTTCTCGGAGCTGAGAGCTATGGAAGAAATCAACCAAGATTGTA 521
OY 534 ACAGGAGACCTGTGTGCTTCTGAAAGAAAGCTGTAGACTGTGAGAGTGGAGAAAGCGAA 593
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 522 ACAGGAGATCTATCAGCTCTCTGAGCAAGAACTATGATTTGATATAGTATATCAAC 581
OY 594 GCTTCTTACATGATGAGC---AGTATCAATCGTTTCAATGGGTTTAAAGATGGTGC 650
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 582 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 641
OY 651 ATTGCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 710
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 642 ATTGACACAGAGAAAGATTAATCTTATCAAGAAAGTATGATGATGATGATGATGATGAT 701
OY 711 ATACAGACAGAGTATCAATTCATGACGATATGAAACTCTAATATGTCAGATGAGATGCA 770
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 702 TTCAATAGAAAGTGTGACATTTGATGATGATGATGATGATGATGATGATGATGATGAT 761
OY 771 GAATCAGAGACAGAGACAGCGTTCTTAAGCGCCATCTTGAGCAACCAATTAATGATG 830
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 762 GCGTTACTAGAAAGCTGTAGCGGCTCAGCCAGTTAGTTG-----GTATCTGT 809
OY 831 ATTGATGCAAAAGATTTTCTATTTATACACCGGGGAAATTTATGATGAGAGAACTGACA 890
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 810 GGGAGGAGAGAGCGCTTCAATTAAGGAAATTTCTGAGGAAATTTCTGAGGCAATGTCACA 869
OY 891 AGTCCATATGAGATTAATCACTTGTGTTTACTTGTGGTATAGCTTCAGCGGATGATG 950
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 870 TCAATG-----GACACAGCATGCTCATCTGATGATGATGATGATGATGATGATGAT 920
OY 951 GATTACTGATAGCAAAAAATTCATGAGGAGAGAAATTTGGGAGAGAAATGATTAATTTGG 1010
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 921 GATTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 980
OY 1011 ATCCAAAGAAACCGGCTAATTTATAGAGTGTGGGATGATGATTAATTTGCTTATAC 1070
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 981 ATGACAGCGTACACCGGCAACGCAAGAGATGTGGGATCAACATGCTGCTTATAT 1040
OY 1071 CCAACCAA 1078
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1041 CCCATCAA 1048
```

RESULT 4
US-08-821-994-60

```
; Sequence 60, Application US/08821994A
; Patent No. 6228643
; GENERAL INFORMATION:
; APPLICANT: Greenland, Andrew J
; APPLICANT: Thomas, Didier RP
; APPLICANT: Jepson, Ian
; TITLE OF INVENTION: Promoters
; FILE REFERENCE: PPD 50108
; CURRENT APPLICATION NUMBER: US/08/821,994A
; EARLIER FILING DATE: 1997-03-22
; EARLIER APPLICATION NUMBER: PCT/GB97/00729
; EARLIER FILING DATE: 1997-03-18
; EARLIER APPLICATION NUMBER: GB 9606062.9
; NUMBER OF SEQ ID NOS: 89
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 60
; LENGTH: 1553
; TYPE: DNA
; ORGANISM: Brassica napus
US-08-821-994-60

Query Match      12.7%; Score 147; DB 4; Length 1553;
Best Local Similarity 54.6%; Pred. No. 2.7e-36;
Matches 372; Conservative 0; Mismatches 285; Indels 24; Gaps 3;
```

```
QY 401 TTGTACCATCCACCTGCATCATGGATTGGAGAAAAAGGTGTCATCCACCACTAA 460
DB 397 TGGGGGCAAGTTCAGATTCTGTGATTGGAGAAAGAGCTGTACTAAGTCTAA 456
QY 461 GTACCAAGGGGGCTGTGAAAGGGTTGGCGTTTCTGCCACGAGCAATGAAAGCAGC 520
DB 457 ACATCAAGAGAGCTGGGAGCGGTGGTCTTCTCGCGAGCTGAGCAATGAAAGAAAT 516
QY 521 ACATGCAATAGCAACAGAGAGACCTTGTAGCCTTTCTGAAACAAGAACTGTAGACTGT 580
DB 517 CAACGAGATTGAACAGAGAGATCTCATCAGCCTCTCTGAGCGAAGCAATGATGTGA 576
QY 581 GGAAGAAAGCGAAGTTCTTCAATGATGGC---AGTATCATGTTGGAATGGGTTT 637
DB 577 CAAGTCTTCAACGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 636
QY 638 AGACATGGTGGCATTTGCCACTGATGATGATGATGATGATGATGATGATGATGATG 697
DB 637 TAAAAACCATGGAATCGACACAGAGAAAGATTATCCTTATCAAGAACGATGAGCACTG 696
QY 698 CAAAGCCATTAAGATACAGACAGAGATTTCAATGACGATATGAACCTTAATATGTC 757
DB 697 TAAAAAAGATTGAATGAAGAGATTTGTGACAAATTGATAGCTATGCGGCTAAATAATC 756
QY 758 ACATGAGATACAGATACAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 817
DB 757 AAACGACGAGAAAGCTTACTGGAAGCTGTAGCGGCTGACGCCAGTTAAGTGTG----- 809
QY 818 AATTAGTGTCTCAATTTGATGCAAAAAGATTTTCAATTTATACACCGGGGGAATTTATGATG 877
DB 810 -----GCATCTGTGGCAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 863
QY 878 AAAAAAAGTAAAGTCCGATATGAGATTATATCATCTTTGTTTACTTTGTTGTTATGCTTC 937
DB 864 -----GCCCATGTTCAACATCATTTGACACGACGACGATGCTATGAGGATACGCTTC 915
QY 938 ACCGATGCTGTAGATTACTGATAGAGCAAAATTCATGCGGAGAGATTTGGGAGAGAGA 997
DB 916 ACAGAAAGGCTGTGATTTACTGATGCTGTAAGAACTTTTGGGAAAGAGTTGGGGTATGGA 975
QY 998 TGGTACATTTGGATCCAAAGAAACAGCGGTAAATTTATTAGGAGTGTGCGATGATTA 1057
DB 976 TGGGTTTATGCAATGCAACGATTAACACCGGCAACTGGAAGGATGTGTGATCATATAT 1035
QY 1058 TTTTGGTTCATACCCACCAA 1078
DB 1036 GCTCGCTTCTGTATCCATCAA 1056
```

```
RESULT 5
US-09-325-932A-109
; Sequence 109, Application US/09325932A
; Patent No. 6451604
; GENERAL INFORMATION:
; APPLICANT: Fliin, Barry
; APPLICANT: Lasham, Annette
; TITLE OF INVENTION: Compositions affecting programmed cell
; FILE REFERENCE: 1022
; CURRENT APPLICATION NUMBER: US/09/325,932A
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 206
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 109
; LENGTH: 1062
; TYPE: DNA
; ORGANISM: Eucalyptus grandis
US-09-325-932A-109

Query Match      11.6%; Score 134.2; DB 4; Length 1062;
Best Local Similarity 51.5%; Pred. No. 2.4e-32;
Matches 452; Conservative 0; Mismatches 393; Indels 32; Gaps 5;
```

```
QY 135 CTGTTCCAACTATGGAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 194
DB 194 CTCTTCCGAGATGCGAGGCGAGCAACAAGAGTGTACAAACGACGCGGGAGGCCGAG 253
QY 195 AAGAGACTTGAGATTTTCAAGATTAATCTGAACTATATCAGGAGATGAAATGCAACAGA 254
DB 254 AGCGGCTGGAAGACTTCAAAAGAACTGAGCTACGTATGGA---GAGAAGCCGGAGG 310
QY 255 AATCAACCCATTTCTATGCTTTAGGATTTGAACAAGTTTCTGCAATCATCTCTTAAG 314
DB 311 GATGGCGCAACACACATGCGTGGGCGTGAACAAGTTGCGCACTGAGCAACAGAG 370
QY 315 TTACAGAAAAGTACTTGCACACTCCCAAGGATGTGTGCGACCAATCAATAATGGCCAA 374
DB 371 TTTCAGGACCGCTTACTCTTCCAAAGTGAAGAGTGTGTACCAAGTGGAGGGCCAA 430
QY 375 AAGAAAATG-----AAGAAGGAACAATATTTCTGTGACCATCCACTGCATCA 422
DB 431 AGGAGAGCTTGATGAGAAACAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 490
QY 423 TGGGATTTGAGCAAAAAAGGTGTATCATCCCAAGTAAATACCAAGGGGCTGTGAAAG 482
DB 491 CTGATTTGAGGAGACATACCGCATTTGACCGCGCTGAAGAGACGAGAAATGCGAAGT 550
QY 483 GGTGGGCGTTTTTGTGCCAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 542
DB 551 TGTGGGCGTTTCTTTCGACCGGAGCATATGAAGATATATGCGCTCAAGAGCGGGAC 610
QY 543 CTGTGTAAGCTTTTGAACAAAGCACTGTAGACTGTGTGGAAGAAAGCGAAGTCTTATC 602
DB 611 CTGATAGGCTTTCCGAGAGAAAGGCTGTGAGCTGTGACACACACAGAGAGGCTGCGAC 670
QY 603 AATGATGCAATATCATGTTGCAATGAGGTTTGAACAATGAGGCAATGCCACTGAT 662
DB 671 GCGGCTATACATGATGATGCTGTGATGATGATGATGATGATGATGATGATGATGAT 730
QY 663 GATGATTAATCTTACAGACTTAAGAGGTTAGATCAAAAGCCAAATTAAGTATCAAGAA 722
DB 731 GAAGACTATCCCTACACAGGCTCTTTGGCGAGGTTGATGTGCAAGCTCACCAAGAG 790
QY 723 GTTACATTTGAGGATATGAAACTTAATTAATGTAGATGAGATGAGATGAGATGAGACA 782
DB 791 GAGAAACA-----ACAAGGCGGTCAACATTTGATGATGATGATGATGATGATGATG 841
QY 783 GAGCAAGCGTCTTAAAGCGCATCTTGTGAGCAACCAATTAATGATGATGATGATGAT 836
DB 842 GAGGAGGCGCTTCTGACAGGCTATCCAAACAGCGGATTAAGGCTGCGCATGAGAGGCTCG 901
```


Oy	483	GATTGGGCGTTTTCCTGCCAGCGGACCATTAAGACGACATGCAATTAACCAACAGAGAC	542
Db	550	TGTTGGGCATTTTCAACTATTAAGCTCTTGAGAGGGCATTAAGCTATGTCAAACACAGGAAG	609
Oy	543	CTTGTTAGCCCTTCTGAAACAAGAACTGTAAGTGTGTGGAAAGAAAGCAAGGCTTTCAC	602
Db	610	CTGGTTCTTTATCTCCGAGCAACAGTTGTGTGATTTGCAATTAAGAGAACCCAGGCTGCAC	669
Oy	603	AATGATGGCAGTATCAATCGTTTCGAAATGGGCTTTAGAACATGCTGCGCATCTCCACTGAT	662
Db	670	GGAGGGCGTTATGAGACACGCGCTCTCCAAATATATCATGATTAATATGGTGGATCTGCTGAA	729
Oy	663	GATGATTTATCTTTCACAGAGCTAAAGAGGGTACGTACGAAGCCCAATTAAGTACAAAGCAG	722
Db	730	GATGAGTATCCATTAATCTGCTGAAAGCCAGCGAATGCAATCCCTCCAAAGGTTAAACCGAAT	789
Oy	723	GTTTCAATTTGACGGATTTGAAACTCTAATTAATGTCAGATGACAGTCAACAATACAGACAA	782
Db	790	GCAATAGCAGCACTAATTAATGATGGTTT-----TGAAGATGTTCTGCTAACAT	837
Oy	783	GAGCAAGCGTTCTTAAAGCGGCATCTTGAGCAACCAATTAAGTCTCAATTAATGACAAA	842
Db	838	GAAAAAGCTCTTAAGAAGACAGTGGCGCACCAACCTGTGTCTGTGCCATTTAAGCAAGT	897
Oy	843	GATTTTCATTTATTAACACCGGGGGGAATTTATGATGAGAAAACTGACAGTCCGATGGG	902
Db	898	GGTAAAGATTTTCAATTTTAATCAAA--AAGSAGTATTCACATGCTGAATGTGGCACTGAA	954
Oy	903	ATTATACACTTGTGTTTACTGTTGGGTTAAG---TTCAAGCGATGCTTAATTAATCTGG	959
Db	955	CTTGACCAATGAGAGTTGAGCGCGTTGGTTAAGGCAGAGTCTCTGAGGGAATTAATTATGG	1014
Oy	960	ATACGAAAAAATTCATGGGGGGAAGAATGGGGGAGAGATGGTTACTATTGATTCGCAAGA	1019
Db	1015	ATAATGAGAACCTCTTGGGACCTCGAATGGGGAAGAAAGGATATTAATAAATGCAACGA	1074
Oy	1020	AACACGGGTATTTATTAGAGAGTGTGGGATGAATTAATTGCTTCATATCCCAACCA	1078
Db	1075	GATATTGAGACAGTAGAAGGGAAGTGTGCTATTGCCATGCAAGCTTCATATCCGACGAA	1133
RESULT 8			
US-08-821-994-62			
Sequence 62, Application US/08821994A			
Patent No. 628643			
GENERAL INFORMATION:			
APPLICANT: Greenland, Andrew J			
APPLICANT: Thomas, Didier RP			
APPLICANT: Jepson, Ian			
TITLE OF INVENTION: Promoters			
FILE REFERENCE: PPD 50108			
CURRENT APPLICATION NUMBER: US/08/821,994A			
CURRENT FILING DATE: 1997-03-22			
EARLIER APPLICATION NUMBER: PCT/GB97/00729			
EARLIER FILING DATE: 1997-03-18			
EARLIER APPLICATION NUMBER: GB 9606062.9			
EARLIER FILING DATE: 1996-03-22			
NUMBER OF SEQ ID NOS: 89			
SOFTWARE: PatentIn Ver. 2.1			
SEQ ID NO 62			
LENGTH: 1434			
TYPE: DNA			
ORGANISM: Brassica napus			
US-08-821-994-62			

	Query Match	10.4%	Score 120.2:	DB 4:	Length 1434;
	Best Local Similarity	51.8%:	Pred.	No. 7.4e-26;	
	Matches 477:	Conservative	0:	Mismatches 393;	Indels 51; Gaps 7;
OY	170 CTACCATTAACCAAGAGAAGGCAAAAGACTTGTAGATTTTCAGAAATACTCGAACA	TGA	229		
b	233 CACGGCTATTATCACCACCAACAAGACAGAAAAGATTCAATTTTTCCAANAGAACCTTAGATT	G	292		

QY	230	TATCAGGCACATGATATGC	AAACAGAAATATACCCCATTTCTACGTTAGATTAAACA	289
Db	293	CATCCATTTACACACAGAAACA	CA - - GAACGGTACTTACAAAGCTTGGCTTAACCAT	349
QY	290	GTTTTCTGACATCTCTCCTCA	AGAGTTGCACCAAAAAGTACTTGCAAGCTCCCAAGATGT	349
Db	350	ATTGGCTATCTACTACATACGAT	TAGTACCGGAACTTTATACCTCGGGCAAGAACCGACCC	409
QY	350	GTCGCAGCAATCAAAATGGC	CAACAAGAAATGAAAGAAACAAATATTTCTTGAGAC - -	407
Db	410	TGTCCGCGCATCTACTAAGG	CCAAAGAAACGTTAACATGAATAATCTACGCCCGCTAAACGA	469
QY	408	- - - - CATCCACCTCATCATG	AGTTGGATTGGAGAAAAAGAGTGCATATCCACCAAGTAAAGTA	463
Db	470	CGTGGAGGTTCCGGAGACAG	GGTTGACTGTGAGAAAAAGAAAGACCCGTTAATGTCATTAAGA	529
QY	464	CCAAAGGCGGCTGTGGAAG	GGGTGGCGTTTCTGCGACGGGAGCCCATAGAACAGACACA	523
Db	530	CCAAAGAACCTTGGCAAGAT	GTGTTGGCGGTTTTCACACGCTGCAGCGAGTGAAGAGTATPAA	589
QY	524	TGCATTTACCAACAGACAC	CTGTGTAACCTTCTGTAACCAAGAAACCTGTAGACTGTGCGA	583
Db	590	CAGATGCTATACAGAGAC	ACTCGATCTTTGTCCGAACAGAACTGTGACTGTGCAGACA	649
QY	584	A - - - CAAAGCCAAAGCTT	CTTACATGATGATGGACAGTATCATGTTGCAATGGGTTTTAGA	640
Db	650	ATCGTACAAACCAAGCTG	TAAACGGCGGCTTAATGATATTATGCTTTCAATTCATAATGAA	709
QY	641	ACATGTGTGGCATTTGCAC	TGATGATTATCTTTCACAGACGCTAAAGAGGATGATGCA	700
Db	710	AAACGGCGGATTTAAACAC	CGGAAAGACTTTCCTTACACAGGACCAATGCGAAATGCGAA	769
QY	701	AGC - - - CAAATAGATTAA	CAAGCAAGGTTTCAATTTGACGATATGAACACTATATATGTC	757
Db	770	CTCTTACTTAAGATTCAG	AGATTGTAACTATGATATGACATAGAAAGATGTTCCTAGTAA	829
QY	758	AGATGACAGTACAGATAC	AGACAGACAGCGTTCTTAAGCGCCATCTTTGAGACATCC	817
Db	830	AGATGAAA - - - - -	CCGGGTTGAAGAGAGCGAGTTTTCATACACACC	868
QY	818	AATTGTGTCTCAATTGATGC	- - - - - AAAAGATTTTCATTTATACACGGGGAATTTA	871
Db	869	TGTGATGTTCTCTATGAT	GTGCTGATGAGAGCTTTCCACATATACCAATTCGGAATCTT	928
QY	872	TGATGAGAAAACTGTAC	ACAGTCCGATGGAATTAATCATCTTGTGTTTACTTGTGGGTTA	931
Db	929	CACGTGAAAGTGTGTACGAT	- - - - - ATGATGTCACGGTGTGTGTGGCGGTTGTGTTA	979
QY	932	TGGTTCCACCGGATGTTG	ATGACTGATAGGCAAAAATTAATGCGGAGGAATGTTGGGG	991
Db	980	TGGGTCAAGAACCGCGGT	GTGCTATTTGATTTGATTTAGCTACTCTTGGGGTACACGTTGGGG	10399
QY	992	AGAAAGTGTTCATTTGAT	TCCAAAGAAACACGGGTAATTTATTAAGAGTGTGTGGGAT	10511
Db	1040	AGAGATGTGTTACATTA	TGATGATGAGAGAAACGTGGCGCTTAATATCCGCGTAAGTGTGGAT	10999
QY	1052	GAATTATTTGCTTCAT	ATACC 1072	
Db	1100	TGCGATAGAACCTTCGAT	CC 1120	

```

RESULT 9
US-08-821-994-61
: Sequence 61, Application US/08821994A
: Patent No. 6228643
:
GENERAL INFORMATION:
: APPLICANT: Greenland, Andrew J
: APPLICANT: Thomas, Didier RP
: APPLICANT: Jepson, Ian
: TITLE OF INVENTION: Promoters
: FILE REFERENCE: PPD 50108
: CURRENT APPLICATION NUMBER: US/08/821,994A

```

```
; CURRENT FILING DATE: 1997-03-22
; EARLIER APPLICATION NUMBER: PCT/GB97/00729
; EARLIER FILING DATE: 1997-03-18
; EARLIER APPLICATION NUMBER: GB 9606062.9
; EARLIER FILING DATE: 1996-03-22
; NUMBER OF SEQ ID NOS: 89
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 61
; LENGTH: 1390
; TYPE: DNA
; ORGANISM: Brassica napus
US-08-821-994-61
```

```
Query Match      10.3%; Score 118.6; DB 4; Length 1390;
Best Local Similarity 51.7%; Pred. No. 2,3e-27;
Matches 476; Conservative 0; Mismatches 394; Indels 51; Gaps 7;
```

```
QY 170 CTACCATTAACCGAGAGAGAGCGCAAGAGAGCTTGAGATTTTCAGAAATTAACGAACTA 229
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 233 CAACGGTATTATCAACCAAGATGAAGATTCATATTTTCAAGAACCAACCTAAGATT 292
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 230 TATCAGGACATGAATGCCAAACAAATACACCCCATTCATCGTTTAGATGAACTA 289
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 293 CATGATCTACACAGAGAGAACAA---GAACGCTACTTACAGCTTGCTTACCAT 349
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 290 GTTTGCTGACATCACTCCTCAGAGTTCCAGAAAAGTACTTGCAGCTCCCAAGATGT 349
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 350 ATTGCTGATCTCTACTACGATGAGTACCGAGATTATACCTGGGGCAAGAACCGAGCC 409
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 350 GTCGACCAATCAAAATGGCCAAAGAAATGAGAGAGACATATTTCTTGAC-- 407
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 410 TGTCCGCCGATCACTAAGCCAGAACGTTAACATGAATATCTACGCCGAGTAAACGA 469
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 408 ----CATCCACCTGCATCATGGGATTTGGAGAAAAAGGTGTCTACCCCAAGTAAAGTA 463
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 470 CGTGAGAGTTCCGGAGAGCGTTGACTGAGACAGAGAGAGAGCCCTTATGCAATTAATA 529
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 464 CCAAGGGGCTGTGAGAGGGGTTGGGCTTTTTCGCCAGGAGCCATNAGAGAGAGACA 523
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 530 CCAAGATCTTGGGAGAGTTGTGGCGTTTTCACACAGCTGCAGCAGTACAGAGCATRAA 589
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 524 TGCATTAAGCAGAGAGACCTTGTAGCCTTTCGAACAAAGCACTGTAGACTGTGGA 583
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 590 CAAATGTAAGCAGAGAGAGCTCATATCTGTGTCGAACAAAGCACTGTAGACTGTGGA 649
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 584 ---AGAAAGGAGAGCTTTCATGATGATGACATATCATCTGCAATGGGTTTGA 640
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 650 ATCTATACCAAGCGGTGTAAGCGGCTTAATGATTAATGCTTTTCAATTCATCATGAA 709
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 641 ACATGCTGATTCGCACTGATGATTAATCCTTACAGAGCTTAAGAGGTAGATGAA 700
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 710 AAACGGGGATTAACACCGAGCAAGACTATCTTACACAGGAAACCAATGGCAATGCAA 769
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 701 AGCCAAATAGATACAGACAGGTT---ACAAATGACGATATGAACCTTAATATGTC 757
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 770 CTCTTACTTAAATAATTCGAGAGTTGTGACTATGATGATGACAAAGATGTTCTAGTAA 829
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 758 AGATGAGATACAGAAATCAGAGAGAGAGCTTCTTAAGCCCATCTTGTAGCAACC 817
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 830 AGATGAAA-----CCGCTTGAAGAGACAGTGTGTCACGACC 868
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 818 AATTAAGTCTCATTTATGATG---AAAAGATTTTCATTTATACCGGGGAAATTTA 871
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 869 TGTAGTGTCTATTTGATGTGTGGAAGAGCTTTCACAACTTACCAATGTGAAATCTT 928
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 872 TGATGAGAAAACTGTACAACTCGTATGAGGATTAATCACTTTGTTTACTTGGGTTA 931
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 929 CACTGAAAGTGTGTAGACT-----ATGATACAGCTGTGTGGGGTTGTTA 979
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 932 TGTGTTCAGGATGTGTAGATTACTGTGATGCGAAAAATTCATGTGGGAGAGATGGGG 991
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 980 TGTATCAGAGAAAGGTGTGTACTATGTGATGTAGTAACTCTTGGGGTACAACTGGGG 1039
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
QY 992 AGAAGATGTTACATTTGGATCCAAAGAAACAGCGTAATTTATTAGAGCTGTGGGAT 1051
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1040 AGAGATGTTTACATTTAGATGAGAGAAACGTGGCGTCCAAATCCGTAAGTGGGAT 1099
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1052 GAATTAATTCGCTTCATACC 1072
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1100 TGGCATTGAAGCCCTGTATCC 1120
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

RESULT 10

```
US-08-821-994-63
; Sequence 63, Application US/08821994A
; Patent No. 6228643
; GENERAL INFORMATION:
; APPLICANT: Greenland, Andrew J
; APPLICANT: Thomas, Didier RP
; APPLICANT: Jepson, Ian
; TITLE OF INVENTION: Promoters
; FILE REFERENCE: PPD 50108
; CURRENT APPLICATION NUMBER: US/08/821,994A
; CURRENT FILING DATE: 1997-03-22
; EARLIER APPLICATION NUMBER: PCT/GB97/00729
; EARLIER FILING DATE: 1997-03-18
; EARLIER APPLICATION NUMBER: GB 9606062.9
; NUMBER OF SEQ ID NOS: 89
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 63
; LENGTH: 1441
; TYPE: DNA
; ORGANISM: Brassica napus
US-08-821-994-63
```

```
Query Match      10.3%; Score 118.6; DB 4; Length 1441;
Best Local Similarity 51.7%; Pred. No. 2,4e-27;
Matches 476; Conservative 0; Mismatches 394; Indels 51; Gaps 7;
```

```
QY 170 CTACCATTAACCGAGAGAGAGCGCAAGAGAGCTTGAGATTTTCAGAAATTAACGAACTA 229
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 219 CAACGGTATTATCAACCAAGATGAAGATTAATATTTTCAAGAACCAACCTAAGATT 278
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 230 TATCAGGAGACATGAATGCACAAACAGAAATACACCCATTCATGTTTAGATGAACAA 289
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 279 CATGCTGATCTACACAGAGAACAA---GAACGCTACTTACAGCTTGCTTACCAT 335
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 290 GTTTGCTGACATCACTCCTCAGAGTTTACGCAAAAAGTACTTGAACCTCCCAAGATGT 349
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 336 ATTGCTGATCTCACTACGATGATGACCGAGTTTATACCTGGGGCAAGAACCGAGCC 395
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 350 GTCGACCAATCAAAATGGCCAAACAGAAATGAGAGAGACATATTTCTTGTGAC-- 407
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 396 TGTCCGCCGATCACTAAGCCAGCAAGAACGTTAATGATTAATCTGAGCCAGTAAACGA 455
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 350 GTCGACCAATCAAAATGGCCAAACAGAAATGAGAGAGACATATTTCTTGTGAC-- 407
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 396 TGTCCGCCGATCACTAAGCCAGCAAGAACGTTAATGATTAATCTGAGCCAGTAAACGA 455
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 408 ----CATCCACCTGCATCATGGGATTTGGAGAAAAAGGTGTCTATCAACCCCAAGTAAAGTA 463
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 456 CGTGAGAGTTCCGGAGAGCGTTGACTGAGAGACAGAGAGAGCCCTTATGCAATTAATA 515
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 464 CCAAGGGGCTGTGGAAGGGGTTGGGCTTTTCGCCAGGAGCCATAGAGAGAGACA 523
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 516 CCAAGATCTTGGCGAAGTGTGTGGGCTTTTCAACAGCTGCACCACTAGAGAGCATRAA 575
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 524 TGCATTAAGCAGAGAGACCTTGTAGCCTTTCGAACAAAGCACTGTAGACTGTGGA 583
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 576 CAAATGTAAGCAGAGAGAGCTCATATCTGTCTCGAAGCAAGAACTTGTGCACTCGCAA 635
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 584 ---AGAAAGGAGAGCTTTCATGATGAGAGAGCACTATCAATCGTCAATGGGTTTGA 640
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 636 ATCATTAACCAAGCGTGAACGCGGCTTAATGATTAATGCTTTTCAATTCATCATGAA 695
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 641 ACATGCTGATTCGCACTGATGATGATTTCTTACAGAGCTTAAGAGGGTATGATGAA 700
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 696 AAACGGGGATTAACACCGAGCAAGACTATCTTACACAGGAAACCAATGGCAATGCAA 755
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```


THIS PAGE BLANK (USPTO)

OY 581 GGAA--GAAAGCGAAGGTTCTTACAAATGANTGGCAGATATCAATTCGTCGAATGGGTTTT 637
 Db 582 CACATCTTATACCAACGAGATGTATGTGGTTTAAATGACTATCAGCTGGTTGATTATAT 641
 OY 638 AGAATCATGTGGCATTTGGCCATGATGATGATTTATCTTACAGACTAAAGAGGGTAGATG 697
 Db 642 CAAATAATGTGTATGTATGTATACCGAAGCGGATTTATCCTTACAAGGCTGCTATGGTCTTG 701
 OY 698 CAAAGCCATAAGATATCAAGACAAAGGTTACAAATGACGGATATGAACTTAATATGTC 757
 Db 702 TGACACCAACAGGAAAAATGCTAAGGTT-----GTTACCATTTGATTC 743
 OY 758 AGATAGAGATACAGAAATACAGACAGACAGACAGCGTTCTTAAGCCCATCCTTGACCAACC 817
 Db 744 ATATATAGATATGTCCTCGTAGAACACGCGAGCGCTCTCTTAAGAAAGCTTTGGCTCACCAACC 803
 OY 818 AATTAGTGTCTCAATTGATGCAAAAGATTTTTCATTTATATACACCGGCGGAATTTATGATGG 877
 Db 804 TATTAGTGTGCGCATTTGATGAAGCGGGTGTGCGGCTTTCACGCTC---TATTCTTGCGGTGT 860
 OY 878 AGAAAACGTACAGATCCGATATGGGATTAATCACTTTGTTTACTGTGGGTATGCTTC 937
 Db 861 ATTTATGAGACTTTGTGGAGCGGAGTATGACACACGCTGTGTGGCTGTGGGTATGAGAAC 920
 OY 938 AGCGGATGTTGTAGTTACTGTGATAGCCAAAATTCATGCGGAGGAGAAAGTTGGCGAAGA 997
 Db 921 CGAAGACGGAAGGAGCTACTGATTTGTAGAAATCTCATGTTGGGAAACAGATGGCGAGAG 980
 OY 998 TGGTTACATTTTGGATCCAAAGAACACCGGTAAATTTATTAGAGAGTGTGGATGAATTA 1057
 Db 981 CGGATTCATTAAGATATGGCAGCGTATTTGAAGCGGCACAGGAAGTGTGGAATCGCTAT 1040
 OY 1058 TTTCGCTTCATACCCACCAAGAGG 1083
 Db 1041 GGAGGCTTCGTATCCGATCAAGAAAG 1066
 RESULT 2
 US-09-938-842A-2449
 : Sequence 2449, Application US/09938842A
 : Patent No. US20020160378A1
 : GENERAL INFORMATION:
 : APPLICANT: Harper, Jeff
 : APPLICANT: Kreps, Joel
 : APPLICANT: Wang, Xun
 : APPLICANT: Zhu, Tong
 : TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
 : TITLE OF INVENTION: SAME AND METHODS OF USE
 : FILE REFERENCE: SCRIPT300-3
 : CURRENT APPLICATION NUMBER: US/09/938, 842A
 : CURRENT FILING DATE: 2001-08-24
 : PRIOR APPLICATION NUMBER: US 60/227, 866
 : PRIOR FILING DATE: 2000-08-24
 : PRIOR APPLICATION NUMBER: US 60/264, 647
 : PRIOR FILING DATE: 2001-01-16
 : PRIOR APPLICATION NUMBER: US 60/300, 111
 : PRIOR FILING DATE: 2001-06-22
 : NUMBER OF SEQ ID NOS: 5379
 : SEQ ID NO 2449
 : LENGTH: 1116
 : TYPE: DNA
 : ORGANISM: Arabidopsis thaliana
 US-09-938-842A-2449

	Query Match	Similarity	Score	DB %	Length
Best Local	471	Conservative	49.68	Pred. No. 1.1e-27	1116
Matches	471	Conservative	0	Mismatches 439	Indels 39
					Gaps 4
0y	136	TTTTCCAAATATGAGAGAGTGCAGCATGCAGCGTGTCTACCATACACGAGAGAGGCA	195		
Db	164	TTTTCCAGTCATGATGATGTCACAAACATGGGAAATGTATACCATCTCGTTCGAGAGAGAAC	223		

OY	196	AGAGCTCTGAGTATTTTCACAAATATCCGACATATCTGAGGGACATGTAAATCGAAACGAA	255
Db	224	GCGCTTTACTATTTTGGAGACACCTCCGTTTATCTACTAACCGGAGCCTGAGA---	280
OY	256	AATCACCCCATCTCATCGTTTAGATTTGACAAAGTTTGCTGACATCACTCCTCAAGAT	315
Db	281	-----ATCTCAGTTACC GGCTTGTTGTAATGGTTTCCGATTATATCTCCATCAAT	334
OY	316	TCAGCAAAAAGTACTTCCAGCTCCCAAGATGTGTGAGAGCAATTCAAAATGGCCACA	375
Db	335	ACGGAGAAATTTGGCATGGGGGTGATCCAAAGCTTCCTAGGAACCAACGCTTCATGACTA	394
OY	376	AGAAATATGAAGAAAGCAATATTTCTTGACCATCCACTCATCATGAGGATGAGAGA	435
Db	395	GCAGCAACCCGATTCACACATGATGTGTATCTTCTCTAATGCCGTTACTGAGAA	454
OY	436	AAAAAGGTGTATCACCACAGTAAAGTACCAGGGGGCTGTGAAAGGGTGTGGCGTTT	495
Db	455	ACGAAAGGGCAGTGCCTGAAAGTCAAAAGTCAACGGCTTTGCGAGGATGTGGCGCTTCT	514
OY	496	CTGGCAGGGGCGCATGGAAGCAGCATGTCATPACACAGSAGACCTGTGTACCTT	555
Db	515	CCACTGTGGAGCAGTGGAAAGCTTAACACAGATTGTGCTGAGAGCTAGTAACTTGT	574
OY	556	CTGAACAGAACTGTGAGTGTGTGGAAAGCAGAGGTTCTTACATGAGTGCAGT	615
Db	575	CTGAGCAAGATTATCATTAATTTGAACAAAGAAACATGTTGCGGAGAGGCAAAATCG	634
OY	616	ATCAATCGTGGATGGGTTTGAACATGTTGGCATTTGCCACTATGATGATTAATCCT	675
Db	635	AGACGCGCTATGATTCATATCATGACAAATGTGTGTGTACCGACACGATTTATCCTT	694
OY	676	ACAGAGCTAAAGAGGTAGATGCAAAAGCCATTAAGTATACAAACAGGTTACAAATGACG	735
Db	695	ACAAAGCTCTCATATGGAGTCTGGGAAGCGCGCTCAAGAGACAA-----	740
OY	736	CATATGAATCTATATATGTCAGATGAGATACAGATCAAGACAGACAGACGCTTCT	795
Db	741	-CAGAAATGTTATATTGATGGGATATGAGAAATTTGCCCTCAAAACATGAAAGCCGCTCA	799
OY	796	TAAAGCGCATCCTTGAGCAACCAATTAGTCTCATATTATGCA-----AAGATTTTC	849
Db	800	TGAAAGCGGTGTCTACACAGCCTGTACTGTCCGTTGTCTATTTCCAGACCGAGAGTTTC	859
OY	850	ATTATATACCGCGGAATTTATGATGAGAAAACTGTACAAGTCCGATATGAGATTAAATC	909
Db	860	AGCTTTATGATCCGGAGTGTGTGACGAACTTGGCGAACAAC-----TAAAC	910
OY	910	ACTTTGTTTACTTTGTGGTTATGTTTCAGCGGATGSGTGAATTAATGATAGAGGAAA	969
Db	911	ATGGGTGTGTGTGTGGGATATGAAACGAGAAATGTGTGTACTGTATGTGATTGGAAA	970
OY	970	ATTCAATGGGGAAGATTGGGGGAGAGATGTTGATTTGGATTCCAAAAGAAACACGGGTA	1029
Db	971	ACTGAGGGGCGACACAAATGGGGGAGGCTGCTACATGAAGATGGCTGCACAACTTGCCA	1030
OY	1030	ATTATTTAGAGTGTGTGGATGATATTATTTGCTTCAATACCAACCA	1078
Db	1031	ATCCAAAGGCGCTATGTGGATGCCAATGCGACCTCATATACCTCTGCA	1079

```

RESULT 3
US-09-938-842A-2484
; Sequence 2484, Application US/09938842A
; Patent No. US20020160378A1
;
; GENERAL INFORMATION:
;
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
;
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; TITLE OF INVENTION: SAME, AND METHODS OF USE
; FILE REFERENCE: SCRIPT300-3

```



```
; TITLE OF INVENTION: SAME, AND METHODS OF USE
; FILE REFERENCE: SRR1300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 742
; LENGTH: 1047
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-742

Query Match      8.9%; Score 102.4; DB 9; Length 1047;
Best Local Similarity 49.7%; Pred. No. 4.9e-20;
Matches 324; Conservative 0; Mismatches 316; Indels 12; Gaps 2;

QY 425 GGAATTGGAGGAAAAAGTGTCATCAACCCCAAGTAAGTACCAGGGGGCTGTGGAAAGGG 484
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 396 GGATTGGAGCAAGAGGGGCTGTACACCCGTTAAGTATCAAGCCGATGTGAGGATG 455
QY 485 TTGGCGGTTTCTGCACAGGAGCCATAGAGCAGCATATAGACAGAGAGACT 544
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 456 CTGGGGCTTTCCGGGCTGGCGGAGTGGAGGATTTACAAAGATTCTAAAGCCAGCT 515
QY 545 TGTAACTTTCTGAACAGAACTCGTAGCTGTG--TGAAGAAAGCAAGGTTCTTA 601
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 516 TGTATGCTATCGAGAACAACTCTTAGCTCGACAGAGACTACAAATAGGATGTG 575
QY 602 CAATGATGCGCATATCAATCGTTCGAATGGTTTGAACATGTCGTCATGCGCACTGA 661
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 576 TGAAGGATATATCGAAGCTTTCGAGTACATATCAAAACCAAGCCATCACTACTGA 635
QY 662 TGTATGTTATCTTAAAGAGCTAAGAGGTAGATGCAAGCCAAATAGATACAGCAA 721
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 636 GGAATCACTATCCATACCAAGATCA-----ACAACTTGTATCTATCACTAC 686
QY 722 GGTATCAATTCAGCATATGAACCTTAATATGTCAGATGAGATGACAGATCAGAGAC 781
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 687 CCTATCATCTCTTTTCGTGACGCCAACAAATCAGCGATATAGAGACGTTCCATGAATA 746
QY 782 AGAGCAAGCTTCTTAAAGCCCATCTGAGCAACCAATAGTGTCAATTTGATGCAAA 841
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 747 TGAAGAAAGCTTCTACAAAGCACTGTCTCAACAGCTGTTTGTGGGATGAGAGGAC 806
QY 842 AGATTTTCATTTATACACCGGGGAAATTTATGTAGAGAAACTGTACAAATCGATG 901
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 807 AGAGGCGCGCTTAAAGCACTACTCGGAGGGGATTCACAGGAAATGTGGAGGATTT 866
QY 902 GATTATCACTTTGTTTACTTGTGGTATGTGTTACGCGGATGTGTATTAATTAATGAT 961
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 867 GCATCATGCGCTTACAAATCTGTTATGGAATGAGTGAAGAAAGCACTAAGTATGGGT 926
QY 962 AGGAAATTTCTAGGGGAGATGAGGAGAAATGTTTACATTTGATTCGCAAGAA 1021
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 927 GGTGAAGAAATTTCTGGGAGAACTTGGGAGAAATGTTTACATGAGTGAAGAGAGA 986
QY 1022 CACGGGTAATTTATAGAGTGTGTGGATGAATTTATTCGCTTCAACCA 1073
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 987 CGTGAATGCAACCCCAAGGATGTGTGTTGGCATTTCTTCTATCA 1038
```

RESULT 6
US-09-887-576-668

; Sequence 668, Application US/09887576
; Patent No. US20020144047A1
; GENERAL INFORMATION:
; APPLICANT: Budworth, P.
; APPLICANT: Brown, D.

```
; APPLICANT: Chang, H.
; APPLICANT: Zhu, T.
; APPLICANT: Han, B.
; APPLICANT: Wang, X.
; APPLICANT: Cooper, Bret
; TITLE OF INVENTION: Promoters for regulation of plant expression
; FILE REFERENCE: 1360.001US1
; CURRENT APPLICATION NUMBER: US/09/887,576
; CURRENT FILING DATE: 2001-06-25
; PRIOR APPLICATION NUMBER: US 60/213,848
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 60/214,087
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 60/258,692
; PRIOR FILING DATE: 2000-12-29
; NUMBER OF SEQ ID NOS: 875
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 668
; LENGTH: 1047
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-887-576-668

Query Match      8.9%; Score 102.4; DB 10; Length 1047;
Best Local Similarity 49.7%; Pred. No. 4.9e-20;
Matches 324; Conservative 0; Mismatches 316; Indels 12; Gaps 2;

QY 425 GGAATTGGAGGAAAAAGTGTCATCAACCCCAAGTAAGTACCAGGGGGCTGTGGAAAGGG 484
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 396 GGATTGGAGCAAGAGGGGCTGTACACCCGTTAAGTATCAAGCCGATGTGAGGATG 455
QY 485 TTGGCGGTTTCTGCACAGGAGCCATAGAGCAGCATATAGACAGAGAGACT 544
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 456 CTGGGGCTTTCCGGGCTGGCGGAGTGGAGGATTTACAAAGATTCTAAAGCCAGCT 515
QY 545 TGTAACTTTCTGAACAGAACTCGTAGCTGTG--TGAAGAAAGCAAGGTTCTTA 601
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 516 TGTATGCTATCGAGAACAACTCTTAGCTCGACAGAGACTACAAATAGGATGTG 575
QY 602 CAATGATGCGCATATCAATCGTTCGAATGGTTTGAACATGTCGTCATGCGCACTGA 661
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 576 TGAAGGATATATCGAAGCTTTCGAGTACATATCAAAACCAAGCCATCACTACTGA 635
QY 662 TGTATGTTATCTTAAAGAGCTAAGAGGTAGATGCAAGCCAAATAGATACAGCAA 721
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 636 GGAATCACTATCCATACCAAGATCA-----ACAACTTGTATCTATCACTAC 686
QY 722 GGTATCAATTCAGCATATGAACCTTAATATGTCAGATGAGATGACAGATCAGAGAC 781
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 687 CCTATCATCTCTTTTCGTGACGCCAACAAATCAGCGATATAGAGACGTTCCATGAATA 746
QY 782 AGAGCAAGCTTCTTAAAGCCCATCTGAGCAACCAATAGTGTCAATTTGATGCAAA 841
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 747 TGAAGAAAGCTTCTACAAAGCACTGTCTCAACAGCTGTTTGTGGGATGAGAGGAC 806
QY 842 AGATTTTCATTTATACACCGGGGAAATTTATGTAGAGAAACTGTACAAATCGATG 901
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 807 AGAGGCGCGCTTAAAGCACTACTCGGAGGGGATTCACAGGAAATGTGGAGGATTT 866
QY 902 GATTATCACTTTGTTTACTTGTGGTATGTGTTACGCGGATGTGTATTAATTAATGAT 961
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 867 GCATCATGCGCTTACAAATCTGTTATGGAATGAGTGAAGAAAGCACTAAGTATGGGT 926
QY 962 AGGAAATTTCTAGGGGAGATGAGGAGAAATGTTTACATTTGATTCGCAAGAA 1021
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 927 GGTGAAGAAATTTCTGGGAGAACTTGGGAGAAATGTTTACATGAGTGAAGAGAGA 986
QY 1022 CACGGGTAATTTATAGAGTGTGTGGATGAATTTATTCGCTTCAACCA 1073
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 987 CGTGAATGCAACCCCAAGGATGTGTGTTGGCATTTCTTCTATCA 1038
```

RESULT 7

US-09-770-445-20/c
: Sequence 20, Application US/09770445
: Patent No. US2002023281A1
: GENERAL INFORMATION:
: APPLICANT: Gorlach, Jorn
: APPLICANT: An, Yong-Qiang
: APPLICANT: Hamilton, Carol M.
: APPLICANT: Price, Jennifer L.
: APPLICANT: Raines, Tracy M.
: APPLICANT: Yu, Yang
: APPLICANT: Rameaka, Joshua G.
: APPLICANT: Page, Amy
: APPLICANT: Mathew, Abraham V.
: APPLICANT: Ledford, Brooke L.
: APPLICANT: Moessner, Jeffrey P.
: APPLICANT: Haas, William David
: APPLICANT: Garcia, Carlos A.
: APPLICANT: Kricker, Maja
: APPLICANT: Slader, Ted
: APPLICANT: Davis, Keith R.
: APPLICANT: Allen, Keith
: APPLICANT: Hoffman, Neil
: APPLICANT: Hurban, Patrick
: TITLE OF INVENTION: Expressed Sequences of Arabidopsis
: FILE REFERENCE: 2023US (PARA-012PRV)
: CURRENT APPLICATION NUMBER: US/09/770,445
: CURRENT FILING DATE: 2001-01-26
: PRIOR APPLICATION NUMBER: US 60/178,472
: PRIOR FILING DATE: 2000-01-27
: NUMBER OF SEQ ID NOS: 999
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 20
: LENGTH: 1282
: TYPE: DNA
: ORGANISM: Arabidopsis thaliana
US-09-770-445-20

Query Match 8.3%; Score 95.8; DB 10; Length 1282;
Best Local Similarity 52.2%; Pred. No. 4.7e-18;
Matches 319; Conservative 0; Mismatches 272; Indels 20; Gaps 4;
QY 414 CCGTCATCATGGATTGGAGAGAAAAAGTGTATCATCCCAAGTAAGTACCAAGGGGC 473
DB 862 CCAGACACAAAAGATTGAGAGAAAGATGTAATTGTCCTGGAAAGAACAGGACAT 803
QY 474 TGTGGAGGGGCTTGGGGCTTTTCTGCCACGGGAGCCATPAGAAGCAGCATGCAATGCA 533
DB 802 TGTGGATCTTGTGGACATTTAGCACAACTGGAGCTCTTGAAGCAGCTTACATCAAGCA 743
QY 534 ACAGGAGACCTGTGTACCTTTGTAACAGAACTGCTAGACTGTG-----AAGAA 587
DB 742 TTGGGAAAAGAAATATCTTTGTCCGAGCAACAGCTTGTGATGTGCTACTTTCAAT 683
QY 588 AGCGAAGTTCTTACATGGATGGACATATCATCTTCAATGGGTTTGAAGCATGT 647
DB 682 AACTTGGTGTGATGGTGGACCTTCTTCAAGCCTTGAATACATTAATACAGGT 623
QY 648 GGCATGCCCTGATGATGATATCTTACAGAGACTAAAGAGGTAGATCAAAAGCAAT 707
DB 622 GGGCTGCACGGAGGAGGCTTATCTTACACCGGAAAGACGGTGGCTCAAAATTTTCA 563
QY 708 AAGATCAAGAACAGGTTACATTTGAGCGATGAACTGAATTAATCTAGATGAGAGT 767
DB 562 GCGAAAACATCGGTGTACAGTCCTGACTCTGTCAACATTAACCTGGGTGCAAGAGAT 503
QY 768 ACAGATCAAGACAGAGAGCGTCTTAAGCGCCATCTTGAAGCAACCAATTAAGTGC 827
DB 502 ---GAACGTGAAGACCGCGGTGGGTGTGAAGGCACTGAGTG-----TGCGGT 456
QY 828 TCATTTGATCAAAAAGTTTTCATTTATACACCGGGGAAATTAATGATGAGAAAGT-G 886
DB 455 TGAGGTGTATCAATCAATTCAGTTTATTAAGAGGAGGATTTTACTAGCAATACATGTGG 396

QY 887 TACAAGTCCGTANAGGATTAATCACTTTGTTTACTTGTGGGTATGTTACAGCGATGG 946
DB 395 TAACACTCCAAATGATGTAAACCATGACAGTGTGAGCTGTTATGAGATTGAAGACGA 336
QY 947 TGTAGATTACTGATACGAAAAATTCATGGGAGAGAGATTGGGAGAGATGGTTACAT 1006
DB 335 CGTCCCTACTGCTTMTAAGAACTCATGGGAGGTGAATGGGAGACATGCTACTT 276
QY 1007 TTGATCCCAA 1017
DB 275 CAAGATGAAA 265
RESULT 8
US-09-938-842A-2351
: Sequence 2351, Application US/09938842A
: Patent No. US20020160378A1
: GENERAL INFORMATION:
: APPLICANT: Harper, Jeff
: APPLICANT: Kreps, Joel
: APPLICANT: Wang, Xun
: APPLICANT: Zhu, Tong
: TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
: FILE REFERENCE: SRRP1300-3
: CURRENT APPLICATION NUMBER: US/09/938,842A
: CURRENT FILING DATE: 2001-08-24
: PRIOR APPLICATION NUMBER: US 60/227,866
: PRIOR FILING DATE: 2000-08-24
: PRIOR APPLICATION NUMBER: US 60/264,647
: PRIOR FILING DATE: 2001-01-16
: PRIOR APPLICATION NUMBER: US 60/300,111
: PRIOR FILING DATE: 2001-06-22
: NUMBER OF SEQ ID NOS: 5379
: SEQ ID NO 2351
: LENGTH: 1095
: TYPE: DNA
: ORGANISM: Arabidopsis thaliana
US-09-938-842A-2351

Query Match 7.5%; Score 86.8; DB 9; Length 1095;
Best Local Similarity 50.6%; Pred. No. 1.9e-15;
Matches 335; Conservative 0; Mismatches 297; Indels 30; Gaps 4;
QY 426 GATTGGAGAAAAAGTGTATCATCCCAAGTAAAGTACCAAGGGGCTGTGAAGGGT 485
DB 391 GATTGGAGAGAAAGAGGCTGTCTAGAGTCAAGATCAACAAGATTGTGAAAGTTGC 450
QY 486 TGGGCGTTTCTGCGACGGGAGCCATGAAAGCAGCATGCAATAGCAACAGAGACCTT 545
DB 451 TGGGCGTTTCTGCGAGTTGCGACAGTGAAGGATTAACAAGATCAGACAAACAACTG 510
QY 546 GTTACCTTTCTGAACAAGAACTGTAGACTGTG---TGAAGAAAGCAAGGTTCTTAC 602
DB 511 GTTTCATTTGCTGAAACAAGAGCTGTGATTTGACACTGAAGAGATCAAGGTTGCA 570
QY 603 AATGATGACGATATCATATGCTTGAATGGGTTTGAACATGGTGCATTTGCCATGAT 662
DB 571 GAGAGTCTCATGAAAGCTGCTTGAATTTATAAAGAAATATGTTGCAATCAAAACCGAA 630
QY 663 GATGATATCTTCTGACAGTAAAGAGGTAGATGCAAAAGCAATAGATCAAGAGAG 722
DB 631 GAGACTTATCTTACGATTTCCAGTGAAGGATTCATATCTGTAGAGCTAATATGTTGGA 690
QY 723 GTTACATTTGACGATATGAAGCTTAATATGTCAAGATGAGATGACAGATCAGAGACA 782
DB 691 GAACCTTAACATCATGATGACAGAGACAGTCCCTGAGATGAT-----TGCGGT 735
QY 783 GAGCAGGTTTCTTAAGCGCATCTTGAAGCAACCAATTAAGTGTCTCAATTTGATGC--- 838
DB 736 GAGGAAAGAACTTCTCAAAAGCTGTGCTACACAGCTGTCTGTAGAGCTAATTTGATGCTGG 795


```

?
?
? TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
?
? TITLE OF INVENTION: SAME, AND METHODS OF USE
?
? FILE REFERENCE: SCRIPI300-3
?
? CURRENT APPLICATION NUMBER: US/09/938, 842A
?
? CURRENT FILING DATE: 2001-08-24
?
? PRIOR APPLICATION NUMBER: US 60/227, 866
?
? PRIOR FILING DATE: 2000-08-24
?
? PRIOR APPLICATION NUMBER: US 60/264, 647
?
? PRIOR FILING DATE: 2001-01-16
?
? PRIOR APPLICATION NUMBER: US 60/300, 111
?
? PRIOR FILING DATE: 2001-06-22
?
? NUMBER OF SEQ ID NOS: 5379
?
? SEQ ID NO 1324
?
? LENGTH: 1122
?
? TYPE: DNA
?
? ORGANISM: Arabidopsis thaliana
?
? US-09-938-842A-1324

```

Query Match	6.48;	Score 74;	DB 9;	Length 1122;
Best Local Similarity	62.48;	Pred. No. 1.1e-11;		
Matches 116; Conservative	0;	Mismatches 70;	Indels 0;	Gaps 0;

QY	407	CCATCCACCTGCATCATGGATTGGAGAAAAAGGTGATCATCACCAGTAAAGTACCA	466
	111	111 111 111111 111111 111111 111111 111111 111111	
Db	414	CGATTTTCCGACGGAGTTCCGATTGGCGTGAACAAAGAGACCTGCACACCTGTTAAANAACA	473
QY	467	AGGGGGCTGTGGAAAGGGGTTGGCGGTTTTCTGCCACGGAGACCCATTAGACAGACACATGC	526
	111111	111111 111111 111111 111111 111111 111111 111111 111111	
Db	474	GGGTATGTGTGGTTCAATGCTGTGGTTCATTATAGTGCATATAGAGACTTGTAAAGAGACACATTT	533
QY	527	AATAGCAACAGAGACCTTGTAGACCTTTCTGAAACAAGAACTCGTAGACGTGTGGAGA	586
	111111	111111 111111 111111 111111 111111 111111 111111 111111	
Db	534	TCTAGCCACTAAAGAGCTGTTAGCCCTACAGACAGACAGCTCGTAGATTGTGTGCATGA	593
QY	587	AAGCGA	592
	111		
Db	594	GTGTGA	599

```

RESULT 12
US-09-878-574-10000
: Sequence 10000, Application US/09878574
: Patent No. US20020110548A1
:
: GENERAL INFORMATION:
: APPLICANT: Byrum, Joseph R.
: APPLICANT: La Rosa, Thomas J.
: APPLICANT: Thompson, Michael D.
: TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
: TITLE OF INVENTION: Plants
: FILE REFERENCE: 38-21(15401)B
: CURRENT APPLICATION NUMBER: US/09/878, 574
: CURRENT FILING DATE: 2001-12-21
: PRIOR APPLICATION NUMBER: 09/333,535
: PRIOR FILING DATE: 1999-06-14
: NUMBER OF SEQ ID NOS: 15775
: SEQ ID NO 10000
: LENGTH: 264
: TYPE: DNA
: ORGANISM: Glycine max
: OTHER INFORMATION: Clone ID: 701103159H1
US-09-878-574-10000

```

Query Match	6.38;	Score 72.4;	DB 10;	Length 264;
Best Local Similarity	66.98;	Pred. No. 1.4e-11;		
Matches 103; Conservative	0;	Mismatches 51;	Indels 0;	Gaps 0;

QY	426	GATTGGAGAGAAAAAGGTGTCATCACCCAACTAAAGTACCAAGG6GGCTGTGGGAAAGGGT	485
Db	64	GATTGGAGACAGAGAGGGGGCAGTTACACCCGTCCAAAGACCAAGCAACTGTGTGGTGCATGT	1233
QY	486	TGGGCGTTTTCGCCACGGAGCCATAGAGCAGCACATGCATATAGCAACAGGAGACCTT	545
Db	124	TGGGCGTTTCAGCTACTAGTGGTGTCATTTCGAAGGCATTAATAATAGATTGTTCACAGGGTCTCTT	1838

QY	546	GTTAGCCTTCTGACACAGA	ACTCGTAGAC	CTGTG	579
Db	184	GTAAGCCTTCTGACACAGA	GCTGATAGAT	TTGTG	217

RESULT 13
US-10-099-275-1

```

? Sequence 1 Application US/10099275
? Publication No. US20020187499A1
? GENERAL INFORMATION:
? APPLICANT: Schneider, Patrick
? APPLICANT: Yamamoto, Karen K.
? APPLICANT: French, Cynthia K.
? APPLICANT: Reprogen, Inc.
? TITLE OF INVENTION: Use of Cathepsin S in the Diagnosis and Treatment of
? TITLE OF INVENTION: Endometriosis
? FILE REFERENCE: 018002-00131005
? CURRENT APPLICATION NUMBER: US/10/099,275
? CURRENT FILING DATE: 2002-04-13
? PRIOR APPLICATION NUMBER: US/09/701,685
? PRIOR FILING DATE: 2001-03-26
? PRIOR APPLICATION NUMBER: US 60/088,017
? PRIOR FILING DATE: 1998-06-04
? NUMBER OF SEQ ID NOS: 4
? SOFTWARE: PatentIn Ver. 2.0
? SEQ ID NO 1
? LENGTH: 1643
? TYPE: DNA
? ORGANISM: Homo sapiens
? FEATURE:
? NAME/KEY: CDS
? LOCATION: (70)..(1065)
? OTHER INFORMATION: human cathepsin S
? US-10-099-275-1

```

Query Match	6.28;	Score 72.2;	DB 9;	Length 1643;
Best Local Similarity	49.28;	Pred. No. 4.4e-11;		
Matches 324;	Conservative 0;	Mismatches 308;	Indels 27;	Gaps 4;

QY	414	CTGTGATCATGAGATTTGAGAGAAAAAGGTGCATCAACCCAGTAAGTACCAAGGGGGC	473
Db	415	CCTGATTTCTGTGGACTGAGAGAGAAAAAGGTGTGTACTGAAAGTAATATCAAGGTTCT	474
QY	474	TGTGGAAAGGGTTGGCGCTTTCTGCCACGGAGCCATAGAAAGCAGACATGCATAGCA	533
Db	475	TGTGTGCTGCTGCTGGGCTTTCAAGTGTGTGGGGCCCTGGAAAGCACAGCTGAAGCTGAAA	534
QY	534	ACAGGAGACCTTTTACGCTTTCTTGAAACAGAACTCGTAGACTGTGGGAAGA-----	588
Db	535	ACAGGAAAGCTGGGTGCTCTAGAGCCACAGAACTGGTGGATGCTCAACTGAAAAATAT	594
QY	587	--AAGCGAAGGTTTCTCAATGAGATGGCAGATATCAATGTTCCAGTGGGTTTTAGAAAT	644
Db	595	GGAAACAAAGGCTGCAATGGTGGCTTATGACAAAGGGTTTCCAGTACATCATTTGATTAAC	655
QY	645	GGTGGCATTTGCCACTGATGATGATTATCTTACAGAGCTAAAGAGGGTAGATGCCAAAGCC	704
Db	655	AAGGCGATCGAGCTAGAGCGCTTCTATTCCTACCAAGAGCA-----TGGATCTGAAAGTGC	709
QY	705	AATAGATATACAGACAAGGTTATCATTTAGACGGAATGAAACTTAAATATGTCAGATGAG	764
Db	710	AATATGACTCAAAATATTCGTGCTGCTGCCACATGTCTCAAGTACACAGCA-ACCTTCCTTAAAGC	768
QY	765	AGTACGAATACAGACACAGACCAAGCGTTCTTAAAGCGCATCTTGACACAACCAATTAGT	824
Db	769	AGAGAAAGTGTGCTGAAAGAAAGCTGTGGCCAAATAAAGC-----CCAGTGTCT	818
QY	825	GTCATCAATGTGCAAAAAGATTTTCATTTATACACGGGGGAATTTATGATGAGCAAAAC	884
Db	817	GTTGGTGTAGATGGCGGCGATCCCTTTCTTCTCTTACACGAAGTGGGTCTACTATGAA	878
QY	885	TGTACAAAGTCCGTATGGGATTAATCACTTTGTGTTTACTGTGGGTTATGGTTCAGCGGAT	944

```
Db 877 CCATCCCTGACTCAGAAATGGAATCATGCTGCTACTTGTGGCTATGAGATCTTAAT 936
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 945 GGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1004
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 937 GGGAAAGAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 996
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1005 ATTGTGATGCAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 1063
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 997 ATTGGATGCGCAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTA 1055
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

RESULT 14

```
US-09-887-576-294
; Sequence 294, Application US/09887576
; Patent No. US20020144047A1
; GENERAL INFORMATION:
; APPLICANT: Budworth, P.
; APPLICANT: Brown, D.
; APPLICANT: Chang, H.
; APPLICANT: Zhu, T.
; APPLICANT: Han, B.
; APPLICANT: Wang, X.
; APPLICANT: Cooper, Bret
; TITLE OF INVENTION: Promoters for regulation of plant expression
; FILE REFERENCE: 1360.001US1
; CURRENT APPLICATION NUMBER: US/09/887,576
; CURRENT FILING DATE: 2001-06-25
; PRIOR APPLICATION NUMBER: US 60/213,848
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 60/214,087
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 60/258,692
; PRIOR FILING DATE: 2000-12-29
; NUMBER OF SEQ ID NOS: 875
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 294
; LENGTH: 2004
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; US-09-887-576-294
```

```
Query Match 6.2%; Score 71.2; DB 10; Length 2004;
Best Local Similarity 58.6%; Pred. No. 9.6e-11;
Matches 143; Conservative 0; Mismatches 98; Indels 3; Gaps 1;
```

```
QY 472 GCTGTGAAAGGGGTGGCGCTTTTCGCCACGGAGCCATAGAAAGCAGCATGCAATAG 531
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1251 GTTTAGGAGCTGTGTGGCATCTCTCAACTATTGGAGCAGTGAAGAAATTAACAAAGATTG 1310
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 532 CACAGAGAGCTTTGAGCTTTCTGAACAAGAACTGAGACTGTGGAG---AAA 588
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1311 TGACAGAGATTTAATATCTTTGTCTGACACAGAAATGGTTGATGTGACACATCTTATA 1370
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 589 GCGAAGCTTCTCAATGATGAGTGCATATCATGCTTGAATGGGTTTGAACATGGTG 648
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1371 ACCAAGATTAATGATGCTGTTAATGACATATGCGTTGATTTATATCAAAATAGTG 1430
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 649 GCATTGCCATGATGATGATTTATCTTACAGAGCTAAAGAGGTAGATGCAAAAGCCATA 708
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1431 GTATTGATCCGAAGCGGATTAATCTTACAAGGCTGCTGATGCTGTTGACACAGAA 1490
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 709 AGAT 712
      | |
Db 1491 GGCT 1494
```

```
RESULT 15
US-09-990-064-1
; Sequence 1, Application US/09990064
; Patent No. US20020164765A1
; GENERAL INFORMATION:
; APPLICANT: Parvinder K. Punia
```

```
; TITLE OF INVENTION: NOVEL COMPOUNDS
; FILE REFERENCE: GP-30242
; CURRENT APPLICATION NUMBER: US/09/990,064
; CURRENT FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: GB0028462.0
; PRIOR FILING DATE: 2000-11-22
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 996
; TYPE: DNA
; ORGANISM: Mammalia
; US-09-990-064-1
```

```
Query Match 5.9%; Score 68.2; DB 9; Length 996;
Best Local Similarity 47.2%; Pred. No. 4.9e-10;
Matches 300; Conservative 0; Mismatches 308; Indels 27; Gaps 2;
```

```
QY 414 CCTGCATCATGGGATGAGGAGAAAGAGTGTCTATCCACCCAGTAAGTAAGTCAAGGGCG 473
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 346 CCTGATTCCTGTGACCTGAGAGAGAGGGGTGTCTCACTGAAGTGAAATATCAGGGTTCT 405
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 474 TGTGGAAGGGGTGGCGCTTTTCGCCACGGAGCCATAGAACAGCACATGCAATAGCA 533
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 406 TGTGTGCTGTGCTGGCGCTTCAGTCCGCTGGGGCTCTGGAAGACACAGTGAAGCTGAAG 465
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 534 ACAGAGACCTTGTGTAGCTTTCTGAAACAAGAACTGCTAGACTGTGT-----GGAA 584
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 466 ACAGGAATCTGTGTCTCTCAGTGCACAGAAATCTGTGATGCTCACTACATCAAGATAC 525
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 585 GAAAGCAAGTCTTTCATGATGAGTGCATATCAATCGTTGCAATGGGTTTGAACAT 644
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 526 GCGAACAAGGCTGCATATGCTGCTTATACAGAGGCGTTCCATATCATCATCGATAC 585
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 645 GTGTGCTTGTGCACTGATGATGATTCCTTACAGAGCTTAAAGAGGTAGATCCAAAGCC 704
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 586 AACGGATCATTCAGAAAGCTTCTTATCCTTACAAAGCCATGATCAAAAGTGCATT-- 643
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 705 AATATGATACAGACAGGTTACATTTGACGATATGAACCTCAATATATGTCAGATGAG 764
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 644 -----ATGACTCAAAACATGCTGTGCTGCACCTGTTCAAAGTACACAGAA 687
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 765 AGTACAGATCAGAGACAGCAAGCGCTTCTTAAAGCCCATCTCTTGACCAACCAATTAGT 824
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 688 CTTCCTTTGGCAGTGAAGAGCGCTTGAAGAGCTGTGSCCAATAAAGACCTGTGTCT 747
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 825 GTCCTCAATTGATGCAAAAGATTTTCATTATATACACCGGGGAAATTTATGATGAGAAAC 884
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 748 GTGGCTTATGATCAAGTCAATCTTCTTCTCTCTTACAGAAAGTGTGTCTATATGAA 807
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 885 TGTACAGTCCGATATGGGATTAATCACTTTGTTTACTTGTGGGTATGTGTACCGGAT 944
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 808 CCATCCCTGACTCAGAAATGGAATCATGAGGCTACTACATTTGGCTATGTAACCTTAAG 867
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 945 GGTGTGATTAATGATGATGAGGAAATTTCAATGGGAGAAATGAGGAGAAATGGTTAC 1004
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 868 GGGAAAGACTACTGCTGTGTGAAAAACAGCTGGGCACTCCACTTGGTGAACAAGATAT 927
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1005 ATTGTGATCCAAGAAACACGGGTAATTTATAGG 1039
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 928 ATTGCGATGCGCAAGAAACAGTAATAAATCATTTGTGG 962
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

```
Search completed: January 11, 2003, 22:08:22
Job time : 70 secs
```